Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Linear Algebra and its Applications 424 (2007) 570–614
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On the state of behaviors

P.A. Fuhrmann [a,∗,1], P. Rapisarda [b], Y. Yamamoto [c]

[a] *Department of Mathematics, Ben-Gurion University of the Negev, Beer Sheva, Israel*
[b] *Information, Signals, Images and Systems (ISIS) Group, School of Electronics and Computer Science, University of Southampton, SO17 1BJ, United Kingdom*
[c] *Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*

Received 22 March 2006; accepted 22 February 2007
Available online 24 March 2007
Submitted by V. Mehrmann

Abstract

The theme of the present paper is the study of the concept of state and the corresponding state maps in the context of Willems' behavioral theory. We concentrate on Markovian system and their representation in terms of first order difference or differential systems. We follow by a full analysis of the special case of state systems, the embedding of a linear system in a state system via the use of state maps arriving at state representations or, equivalently, to a realization theory for behaviors. Minimality is defined and characterized and a state space isomorphism theorem is established. Realization procedures based on the shift realization are developed as well as a rigorous analysis of the construction of state maps. The paper owes much to Rapisarda and Willems [P. Rapisarda, J.C. Willems, State maps for linear systems, SIAM J. Contr. Optim. 35 (1997) 1053–1091].

Keywords: Behaviors; Behavior homomorphisms; Markovian systems; State systems; Realization theory; State maps; State space isomorphisms

∗ Corresponding author.
E-mail address: fuhrmannbgu@gmail.com (P.A. Fuhrmann).
1 Partially supported by ISF under Grant No. 1282/05.

0024-3795/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2007.02.028

1. Introduction

The concept of state is a basic one in systems theory. Nontrivial systems have a memory of past events and the scope of this memory is crystallized in the concept of state. As a byproduct, using the introduced state variables, we are led to first order representations, i.e. realizations, of the system, a construction extremely useful in simulation, control and design applications. The usefulness of first order representations is such that in many cases they are taken as a starting point for the analysis of systems. However, alternative points of view have been adopted in the course of time, the most prominent of which are associated with the names of Kalman, Rosenbrock and Willems. It was Kalman who formalized the input/output approach and put realization theory as a cornerstone of the general linear systems theory. Rosenbrock realized that most linear systems are modelled in terms of higher order equations and, with the introduction of polynomial system matrices, he constructed a beautiful theory that did not take input/output considerations as basic. This was partly hidden because it was a result of noncontrollability or nonobservability which were expressed in terms of noncoprimeness of certain polynomial matrices. It was Willems, in a series of seminal papers who took the final step of disposing with inputs, outputs and 1st order representations, and focused on manifest, or external, variables. Latent, or auxiliary, variables were introduced in this framework in order to accommodate the many practical cases in which, in order to model the behavior of a system, auxiliary dynamics involving additional variables must also be used.

Of course the concept of state with all its usefulness had to be accommodated also in the behavioral setting. The concepts of state systems and state representations were introduced, in the behavioral setting, in Willems [26,27]. There one can find the characterizations of state systems as those having first order representations. This characterization was not constructive. To construct a first order representation for a linear dynamical system, state maps were introduced in Rapisarda and Willems [21].

The aim of this paper is to take another, somewhat different, look at the concepts of state systems, state representations and state maps. The approach taken here to the construction of state maps and 1st order representations, i.e. realizations, of behaviors is based on the theory of polynomial models and its application to behaviors. In particular, we shall employ the characterization of behavior homomorphisms and the analysis of their invertibility properties, as developed in Fuhrmann [10,11].

The paper is structured as follows. In Section 2, we shall collect some preliminary results about polynomial models, the shift realization, and reduction to dual Brunovsky form via output injection. For the analysis of state maps, we review the basic results on behavior homomorphisms and the role played by doubly unimodular embeddings in the analysis of their invertibility properties. Finally, we consider the class of state to output maps studied in Hautus and Heymann [15], and their connection to rational models and autonomous behaviors.

In Section 3 we characterize Markovian systems and their generalization, i.e. l-Markovian systems or equivalently l-memory span systems. We show that the analysis of l-Markovian systems can be reduced to the special case of autonomous system and then to the case of autonomous systems in dual Brunovsky form.

Section 4 is the core of the paper. We begin by analyzing the construction of state maps for autonomous dynamical systems. This analysis is the prototype for the general case. We proceed by showing how to construct state maps for an arbitrary behavior. This uses realizations of autonomous behaviors, doubly coprime factorizations and behavior homomorphisms. We conclude by showing how a special choice of basis, related to the dual Brunovsky form, leads to a simple procedure for the construction of state maps. This recovers results of Rapisarda and Willems [21]. Finally, in Section 5, we illustrate the theory by working out in detail a few examples.

2. Preliminaries

Our interest in this paper is mostly in discrete time systems, therefore we find it unnecessary to restrict ourselves to the real or complex field and we will work with linear spaces over an arbitrary field $\mathbb{F}$. We will begin by giving a concise introduction to polynomial and rational models, first introduced in Fuhrmann [3]. Let $\mathbb{F}$ denote an arbitrary field. We will denote by $\mathbb{F}^m$ the space of all m-vectors with coordinates in $\mathbb{F}$. By $\mathbb{F}((z^{-1}))^m$ we denote the set of vectorial truncated Laurent series, namely, the space of series of the form $g(z) = \sum_{j=-\infty}^{n(g)} g_j z^j$ with $g \in \mathbb{F}^m$ and $n(g) \in \mathbb{Z}$. By $z^{-1}\mathbb{F}[[z^{-1}]]^m$ we denote the subspace of $\mathbb{F}[[z^{-1}]]$ consisting of all formal power series with vanishing constant term.

An infinite sequence $\{x_t\}_{t=1}^{\infty}$, $t \in \mathbb{Z}_+$, $x_t \in \mathbb{F}^n$ is a time trajectory. We associate with it the formal power series $x = \sum_{t=1}^{\infty} x_t z^{-t}$. The space of all time trajectories is $z^{-1}\mathbb{F}[[z^{-1}]]^n$.

The space $\mathbb{F}((z^{-1}))^m$ has the following direct sum decomposition

$$\mathbb{F}((z^{-1}))^m = \mathbb{F}[z]^m \oplus z^{-1}\mathbb{F}[[z^{-1}]]^m \tag{1}$$

and we denote by π_+ and π_- the projections of $\mathbb{F}((z^{-1}))^m$ on $\mathbb{F}[z]^m$ and $z^{-1}\mathbb{F}[[z^{-1}]]^m$ respectively. Clearly, π_+ and π_- are complementary projections. At some point we find it convenient to use row space version of the above spaces. In particular, $\mathbb{F}_r[z]^m$ is the space of m-row vectors with entries in $\mathbb{F}[z]$. The backward shift $\sigma : z^{-1}\mathbb{F}[[z^{-1}]]^n \to z^{-1}\mathbb{F}[[z^{-1}]]^n$ is defined by $\sigma x = \pi_- z x$, or in terms of coordinates, by $(\sigma x)_t = x_{t+1}$.

The space $\mathbb{F}((z^{-1}))^m$ is endowed with a natural $\mathbb{F}[z]$-module structure, given by multiplication with $\mathbb{F}[z]^m$ as a submodule. In particular, $S : \mathbb{F}((z^{-1}))^m \to \mathbb{F}((z^{-1}))^m$ is defined by $Sf(z) = zf(z)$. As $\mathbb{F}[z]^m$ is a submodule, we can induce a module structure on it by restricting the module structure on $\mathbb{F}((z^{-1}))^m$. In particular, we define $S_+ : \mathbb{F}[z]^m \to \mathbb{F}[z]^m$ by $S_+ = S|\mathbb{F}[z]^m$.

We can induce in the space $z^{-1}\mathbb{F}[[z^{-1}]]^m$ an $\mathbb{F}[z]$-module structure via the isomorphism $z^{-1}\mathbb{F}[[z^{-1}]]^m \simeq \mathbb{F}((z^{-1}))^m/\mathbb{F}[z]^m$. This $\mathbb{F}[z]$-module structure is equal to the one induced by the **left** or **backward shift operator** S_- or, for reasons of compatibility with behavioral theory usage, σ defined by $S_- h = \sigma h = \pi_- z h$, $h \in z^{-1}\mathbb{F}[[z^{-1}]]^m$. Similarly, given a rational function G, we define the **Hankel operator** $H_G : \mathbb{F}[z]^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ by

$$H_G u = \pi_- G u, \quad u \in \mathbb{F}[z]^m. \tag{2}$$

Any $\mathbb{F}[z]$-submodule $\mathscr{M} \subset \mathbb{F}[z]^m$ has a representation $\mathscr{M} = M(z)\mathbb{F}[z]^k$ for some $m \times k$ polynomial matrix. If we require M to have full column rank, then $M(z)$ is uniquely determined up to a right unimodular factor. Given a $p \times m$ polynomial matrix $R(z)$, the set $\mathscr{M} = \{f \in \mathbb{F}[z]^m | R(z)f(z) = 0\}$ is a submodule, hence has a representation $\mathscr{M} = M(z)\mathbb{F}[z]^k$. We call M a **minimal right annihilator**, or **MRA** for short, of R. Similarly, given a $p \times m$ polynomial matrix $R(z)$, we say M is a **minimal left annihilator**, or **MLA** for short, for R if $\widetilde{M}$ is a MRA of $\widetilde{R}$. Here $\widetilde{R}$ denotes the transpose of the polynomial matrix R. Note that a MLA is always left prime.

$\mathbb{F}[z]^p$, besides being an $\mathbb{F}[z]$-module, has also a naturally induced $\mathbb{F}[z^{-1}]^p$-module structure defined by

$$\sigma_+ f = \pi_+ z^{-1} f = \frac{f(z) - f(0)}{z}, \quad f \in \mathbb{F}[z]^p. \tag{3}$$

We will refer to σ_+ as the **downward shift operator**.

Given a nonsingular polynomial matrix D in $\mathbb{F}[z]^{m\times m}$ we define two projections π_D in $\mathbb{F}[z]^m$ and π^D in $z^{-1}\mathbb{F}[[z^{-1}]]^m$ by letting $\pi_D f = D\pi_- D^{-1} f$ for $f \in \mathbb{F}^m[z]$ and $\pi^D h = \pi_- D^{-1}\pi_+ Dh$ for $h \in z^{-1}\mathbb{F}[[z^{-1}]]^m$. We define two linear subspaces of $x_D \subset \mathbb{F}[z]^m$ and $X^D \subset z^{-1}\mathbb{F}[[z^{-1}]]^m$ by $X_D = \operatorname{Im}\pi_D$ and $X^D = \operatorname{Im}\pi^D$. We refer to X_D as a **polynomial model** whereas to X^D as a **rational model**. For the details of polynomial model theory, we refer to Fuhrmann [3,4,6,7,10].

2.1. The shift realization

The following is a version of the shift realization as proved in Fuhrmann [3,4].

Theorem 2.1. *Let $G = VT^{-1}U + W$ be a representation of a proper, $p \times m$ rational function. In the state space X_T a system is defined by*

$$\begin{cases} Af = S_T f & f \in X_T, \\ B\xi = \pi_T U\xi & \xi \in \mathbb{F}^m, \\ Cf = (VT^{-1}f)_{-1} & f \in X_T, \\ D = G(\infty). \end{cases} \tag{4}$$

Then this is a realization of G. This realization is observable if and only if V and T are right coprime and it is reachable if and only if T and U are left coprime. We will call (4) the **shift realization** *and denote it by $\Sigma(VT^{-1}U + W)$.*

We note that no coprimeness assumptions are made as far as the realization itself is concerned. The coprimeness assumptions relate to reachability and observability. In particular, this allows us to realize systems with no inputs or outputs. This of course turns out to be very useful when dealing with the class of finite dimensional behaviors. This class will turn out to be equal to the class of autonomous behaviors, as will be seen in Section 3.

A special case of importance for us is the case of a nonsingular polynomial matrix $T(z)$ considered as a left denominator of a matrix fraction. We define the pair (C_T, A_T), acting in the state space X_T, by

$$\begin{cases} A_T f = S_T f & f \in X_T, \\ C_T f = (T^{-1}f)_{-1} & f \in X_T. \end{cases} \tag{5}$$

Note that in the realization (5) the pair (C_T, A_T) depends only on T. An isomorphic pair is obtain by taking the state space to be the rational model X^T with (C^T, A^T) defined by

$$\begin{cases} A^T f = S^T f & f \in X_T, \\ C^T f = (f)_{-1} & f \in X^T. \end{cases} \tag{6}$$

2.2. Bases and the Brunovsky form

For the analysis of state maps, we are interested in the construction of bases for polynomial and rational models. Given a nonsingular polynomial matrix $D(z) \in \mathbb{F}[z]^{m\times m}$, let $D(z) = D_0 + \cdots + D_s z^s$. We assume D_s is nonzero but no assumption on its nonsingularity is made. We clearly have the direct sum representation $z^{-1}\mathbb{F}[[z^{-1}]]^m = \mathbb{F}_s[z]^m \oplus z^{-s-1}\mathbb{F}[[z^{-1}]]^m$, where $\mathbb{F}_s[z]^m = \{\sum_{i=1}^s \frac{\xi_i}{z^i} | \xi_i \in \mathbb{F}^m\}$. Since, for $h \in z^{-s-1}\mathbb{F}[[z^{-1}]]^m$, Dh is clearly strictly proper, we have $\pi^D h = \pi_- D^{-1}\pi_+ Dh = 0$. Therefore we have

$$X^D = \pi^D \mathbb{F}_s[z]^m = \pi^D \sum_{i=1}^{s} \frac{\xi_i}{z^i} = \sum_{i=1}^{s} \pi_- D^{-1} \pi_+ D(z) z^{-i} \xi_i$$
$$= \sum_{i=1}^{s} \pi_- D^{-1} \pi_+ z^{-i} D(z) \xi_i = \sum_{i=1}^{s} \pi_- D^{-1} E_i(z) \xi_i,$$

where the polynomial matrices E_i are defined by $E_i(z) = \pi_+ z^{-i} D(z)$, for $i = 1, \ldots, s$. In the scalar case the polynomials $E_1, \ldots, E_s$, attributed by Kalman [17] to Tschirnhausen's work as presented in Weber [24], are the basis elements related to the control canonical form, see also Fuhrmann [5,8]. The polynomial matrices E_i are the multivariable generalizations and were introduced in Fuhrmann [5] in the analysis of state feedback. If $D(z)$ is properly invertible, i.e. its inverse is a proper rational matrix, which is the case if $D(z)$ is either row or column proper, then it is easily checked that we have $\pi_D E_j = E_j$.

Actually, in the analysis of state maps for behaviors, we will need an extension of this procedure to rectangular polynomial matrices. Thus, if $R(z)$ is a $p \times m$ polynomial matrix of degree ν, then we define $R_i(z) = \pi_+ z^{-i} R(z)$ for $i = 1, \ldots, \nu$.

Next, suppose the nonsingular polynomial matrix $D(z)$ has the representation

$$D(z) = \mathrm{diag}(z^{\nu_1}, \ldots, z^{\nu_p}). \tag{7}$$

We will refer to (7) as the **polynomial Brunovsky matrix**. If we apply the above procedure to the Brunovsky matrix, we obtain, (after rearranging columns), the **standard basis matrix** defined by

$$H(z) = \left(\begin{array}{ccccc|ccc|ccccc}
1 & z & \cdot & \cdot & z^{\nu_1-1} & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 \\
0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 \\
0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 1 & z & \cdot & \cdot & z^{\nu_p-1}
\end{array}\right). \tag{8}$$

The matrix representation of the pair (C_D, A_D) defined in (5), with respect to this basis, is given by

$$\begin{aligned} A_D &= \mathrm{diag}(N_{\nu_1}, \ldots, N_{\nu_p}), \\ C_D &= \mathrm{diag}(L_{\nu_1}, \ldots, L_{\nu_p}), \end{aligned} \tag{9}$$

where the $1 \times \nu$ and $\nu \times \nu$ matrices L_ν and N_ν are defined by

$$L_\nu = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix}$$

$$N_\nu = \begin{pmatrix} 0 & & & & & 0 \\ 1 & \cdot & & & & \cdot \\ & \cdot & \cdot & & & \cdot \\ & & \cdot & \cdot & & \cdot \\ & & & \cdot & \cdot & 0 \\ & & & & 1 & 0 \end{pmatrix}. \tag{10}$$

2.3. Behaviors and behavior homomorphisms

In $z^{-1}\mathbb{F}[[z^{-1}]]^m$ we define the projections P_n, $n \in \mathbf{Z}_+$ by

$$P_n \sum_{i=1}^{\infty} \frac{h_i}{z^i} = \sum_{i=1}^{n} \frac{h_i}{z^i}. \tag{11}$$

We say that a subset $\mathscr{B} \subset z^{-1}\mathbb{F}[[z^{-1}]]^m$ is **complete** if for any $w = \sum_{i=1}^{\infty} w_i z^{-i} \in z^{-1}\mathbb{F}^m[[z^{-1}]]$ and for each positive integer N, $P_N w \in P_N(\mathscr{B})$ implies $w \in \mathscr{B}$.

A **behavior** in our context is defined as a linear, shift invariant and complete subspace of $z^{-1}\mathbb{F}[[z^{-1}]]^m$. Behaviors can be algebraically characterized. A basic result of behavioral theory, see Willems [25, Theorem 5] or Fuhrmann [10], is that a subspace $\mathscr{B} \subset z^{-1}\mathbb{F}[[z^{-1}]]^m$ is a behavior if and only if it admits a kernel representation of the form $\mathscr{B} = \operatorname{Ker} R(\sigma)$. A special class of interest is that of finite dimensional behaviors. A behavior $\mathscr{B} \subset z^{-1}\mathbb{F}[[z^{-1}]]^m$ is finite dimensional if and only if it is a rational model. To emphasize the connection of behaviors to rational models, we will use also the notation $X^R = \operatorname{Ker} R(\sigma)$. This is justified by Proposition 2.1. A **linear dynamical system** is a triple $(\mathbb{T}, X, \mathscr{B})$, where $\mathbb{T} \subset \mathbb{R}$ is a time set, usually taken to be $\mathbb{R}$ for continuous time systems, and $\mathbb{Z}$ or $\mathbb{Z}_+$ for discrete time systems. The vector space X is identified with $\mathbb{F}^m$ and $\mathscr{B} \subset (\mathbb{F}^m)^{\mathbb{T}}$ is a corresponding behavior.

A central tool in behavior theory, introduced in Fuhrmann [9,10] is that of a behavior homomorphism. Given two behaviors $\mathscr{B}_1, \mathscr{B}_2$, we define for the backward shift operator σ its restriction to the behaviors by $\sigma^{\mathscr{B}_i} = \sigma|\mathscr{B}_i$. If the behaviors are given in kernel representations $\mathscr{B}_i = \operatorname{Ker} P_i(\sigma)$, we will write also σ^{P_i} for $\sigma^{\mathscr{B}_i}$. A behavior homomorphism $Z : \mathscr{B}_1 \to \mathscr{B}_2$ is an $\mathbb{F}[z]$-homomorphism with respect to the natural $\mathbb{F}[z]$-module structure in the behaviors, i.e. it satisfies $Z\sigma^{P_1} = \sigma^{P_2} Z$. Our interest is in the characterization of behavior homomorphisms. It turns out that no general characterization of behavior homomorphisms is available. However, adding some continuity constraints makes the problem tractable by duality theory. The appropriate continuity is with respect to the weak* topologies on the two behaviors. For a full discussion of this see Fuhrmann [11] and, in particular, Theorem 3.4 there which we quote here.

Theorem 2.2. *Let $M \in \mathbb{F}[z]^{p\times m}$ and $\overline{M} \in \mathbb{F}[z]^{\bar{p}\times \bar{m}}$ be of full row rank. Then* $\operatorname{Ker} M(\sigma)$ *is an $\mathbb{F}[z]$-submodule of $z^{-1}\mathbb{F}[[z^{-1}]]^m$ and* $\operatorname{Ker} \overline{M}(\sigma)$ *is an $\mathbb{F}[z]$-submodule of $z^{-1}\mathbb{F}[[z^{-1}]]^{\bar{m}}$. Moreover $Z : \operatorname{Ker} M(\sigma) \to \operatorname{Ker} \overline{M}(\sigma)$ is a continuous behavior homomorphism if and only if there exist $\overline{U} \in \mathbb{F}[z]^{\bar{p}\times p}$ and U in $\mathbb{F}[z]^{\bar{m}\times m}$ such that*

$$\overline{U}(z)M(z) = \overline{M}(z)U(z) \tag{12}$$

and

$$Zh = U(\sigma)h \quad h \in \operatorname{Ker} M(\sigma). \tag{13}$$

The next theorem, see Theorem 3.6 in Fuhrmann [10], summarizes the invertibility properties of continuous behavior homomorphisms.

Theorem 2.3. *Given two full row rank polynomial matrices $M \in \mathbb{F}[z]^{p\times m}$, $\overline{M} \in \mathbb{F}[z]^{\bar{p}\times \bar{m}}$ describing the behaviors $\mathscr{B} = \operatorname{Ker} M(\sigma)$ and $\overline{\mathscr{B}} = \operatorname{Ker} \overline{M}(\sigma)$ respectively. Let $\overline{U}, U$ be appropriately sized polynomial matrices satisfying*

$$\overline{U}(z)M(z) = \overline{M}(z)U(z), \tag{14}$$

and let $Z : \operatorname{Ker} M(\sigma) \to \operatorname{Ker} \overline{M}(\sigma)$ be the continuous behavior homomorphism defined by

$$Zh = U(\sigma)h = \pi_- Uh \quad h \in \operatorname{Ker} M(\sigma). \tag{15}$$

Then

1. *Z is injective if and only if M, U are right coprime.*
2. *Z is surjective if and only if $\overline{U}, \overline{M}$ are left coprime and*

$$\operatorname{Ker} \begin{pmatrix} -\overline{U}(z) & \overline{M}(z) \end{pmatrix} = \operatorname{Im} \begin{pmatrix} M(z) \\ U(z) \end{pmatrix}. \tag{16}$$

3. *Z as defined above is the zero map if and only if, for some appropriately sized polynomial matrix $L(z)$, we have*
$$U(z) = L(z)M(z). \tag{17}$$
4. *Z defined in (15) is invertible if and only if there exists a doubly unimodular embedding*
$$\begin{pmatrix} \overline{X} & -\overline{Y} \\ -\overline{U} & \overline{M} \end{pmatrix}\begin{pmatrix} M & Y \\ U & X \end{pmatrix} = \begin{pmatrix} M & Y \\ U & X \end{pmatrix}\begin{pmatrix} \overline{X} & -\overline{Y} \\ -\overline{U} & \overline{M} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{18}$$
of $(-\overline{U}(z) \quad \overline{M}(z))$ *and* $\begin{pmatrix} M(z) \\ U(z) \end{pmatrix}$.
5. *If Z is invertible, then in terms of the doubly unimodular embedding (18), its inverse Z^{-1} : $\operatorname{Ker} \overline{M}(\sigma) \to \operatorname{Ker} M(\sigma)$ is given by*
$$Z^{-1} = -\overline{Y}(\sigma). \tag{19}$$

An important tool for the analysis of behaviors is the elimination theorem, see Willems [26, Prop. 4.1.c], Kuijper [19], Polderman [20], which gives a procedure for the elimination of latent variables. The present proof is new and uses the analysis of the invertibility of behavior homomorphisms.

Theorem 2.4. *Let a behavior $\mathscr{B}$ be given by the latent variable representation*
$$Q(\sigma)w = P(\sigma)\xi. \tag{20}$$
Let $N(z)$ be a MLA of $P(z)$ and define $R(z) = N(z)Q(z)$. Then

1. *We have the equality*
$$N(z)\begin{pmatrix} Q(z) & -P(z) \end{pmatrix} = R(z)\begin{pmatrix} I & 0 \end{pmatrix}, \tag{21}$$
with
$$\operatorname{Ker}\begin{pmatrix} -N(z) & R(z) \end{pmatrix} = \operatorname{Im}\begin{pmatrix} Q(z) & -P(z) \\ I & 0 \end{pmatrix} \tag{22}$$
holding.
2. *The projection $\pi_w : \operatorname{Ker}\begin{pmatrix} Q(\sigma) & -P(\sigma) \end{pmatrix} \to \operatorname{Ker} R(\sigma)$, defined by $\pi_w\begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} I & 0 \end{pmatrix}\begin{pmatrix} w \\ x \end{pmatrix} = w$ is a surjective behavior homomorphism.*
3. *A kernel representation of $\mathscr{B}$ is given by*
$$\mathscr{B} = \operatorname{Ker} R(\sigma). \tag{23}$$

Proof

1. Equality (21) follows from the definition of $R(z)$ and the fact that $N(z)$ is a MLA of $P(z)$. Clearly $(-N(z) \quad R(z))$ is left prime since $N(z)$, as a MLA of $P(z)$, is.
Note that $\operatorname{Ker} N(z) = \operatorname{Im} P(z)$ implies $\operatorname{Ker}\begin{pmatrix} N(z) & 0 \end{pmatrix} = \operatorname{Im}\begin{pmatrix} 0 & P(z) \\ I & 0 \end{pmatrix}$ and, using the equality
$$\begin{aligned}(-N(z) \quad R(z))&\begin{pmatrix} Q(z) & P(z) \\ I & 0 \end{pmatrix} \\ &= (-N(z) \quad 0)\begin{pmatrix} I & Q(z) \\ 0 & I \end{pmatrix}\begin{pmatrix} I & -Q(z) \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & -P(z) \\ I & 0 \end{pmatrix},\end{aligned}$$
it follows that $\operatorname{Ker}\begin{pmatrix} -N(z) & R(z) \end{pmatrix} = \operatorname{Im}\begin{pmatrix} Q(z) & -P(z) \\ I & 0 \end{pmatrix}$.

2. Applying Theorem 2.3, using equalities (21) and (22), the statement follows.
3. Follows from part 2. $\square$

A useful fact to remember is that, given the coprime factorizations $\overline{D}^{-1}\overline{N} = ND^{-1}$, then a MLA of $\begin{pmatrix} D \\ N \end{pmatrix}$ is given by $\begin{pmatrix} -\overline{N} & \overline{D} \end{pmatrix}$.

2.4. Input/output maps

From the input/output point of view, a **linear system** is a linear map from an input signal space to an output signal space, both spaces being linear spaces over the field $\mathbb{F}$. To make it concrete, we will focus on discrete time systems. Moreover, we always assume that in the remote past all signals were zero. Thus we identify the **input and output signal spaces** with $\mathbb{F}((z^{-1}))^m$ and $\mathbb{F}((z^{-1}))^p$ respectively. In an expansion $w = \sum_{i=-\infty}^{N_w} w_i z^i$, the z^i are considered as time markers and the w_i are the values of the signal w at time $t = i$. The input and output signal spaces carry a natural $\mathbb{F}[z]$-module structure. An **Input/Output map** $\bar{f} : \mathbb{F}((z^{-1}))^m \to \mathbb{F}((z^{-1}))^p$ is **time invariant** if it is a $\mathbb{F}[z]$-module homomorphism. Input/Output maps are representable by **transfer functions**, i.e. $y = \bar{f}(u) = Gu$ where $G \in \mathbb{F}((z^{-1}))^{p\times m}$. The system is **(strictly) causal** if ($\bar{f}(\mathbb{F}[[z^{-1}]]^m) \subset z^{-1}\mathbb{F}[[z^{-1}]]^p$) $\bar{f}(\mathbb{F}[[z^{-1}]]^m) \subset \mathbb{F}[[z^{-1}]]^p$. These conditions are expressible as ($G \in z^{-1}\mathbb{F}[[z^{-1}]]^{p\times m}$) $G \in \mathbb{F}[[z^{-1}]]^{p\times m}$. Finally, a system is finite dimensional, linear, time invariant if there exists a nonzero polynomial $p \in \mathbb{F}[z]$ for which $pG \in \mathbb{F}[z]^{p\times m}$, i.e. the transfer function is **rational**.

For the development of realization theory, we need also the concept of a **restricted Input/Output map**. Given an Input/Output map $\bar{f}$, the restricted Input/Output map $f : \mathbb{F}[z]^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is defined via the commutative diagram

$$\begin{tikzcd}
\mathbb{F}((z^{-1}))^m \arrow[r, "\overline{f}"] & \mathbb{F}((z^{-1}))^p \arrow[d, "\pi_-"] \\
\mathbb{F}[z]^m \arrow[u, "i"] \arrow[r, "f"] & z^{-1}\mathbb{F}[[z^{-1}]]^m
\end{tikzcd}$$

where $i : \mathbb{F}[z]^m \to \mathbb{F}((z^{-1}))^m$ is the natural embedding $i(g) = g$ and π_- is the projection defined in Section 2. If $\bar{f}(u) = Gu$, then clearly $f(u) = H_G u$, where H_G is the Hankel operator defined by (2). Our interest in Hankel operators in this context is due to the fact that they describe best the restricted input/output map and thus provide a key to realization theory.

2.5. State to output maps

We will find it of importance to study in somewhat more detail a special class of systems, namely the state to output (state/output) systems. In this connection, see Hautus and Heymann [15]. In some cases the input of a given system does not play a significant role, as is the case in the study of observers. It is therefore of interest to give characterizations of the transfer functions of such special systems. Thus a state to output transfer function is representable as $G(z) = C(zI - A)^{-1}$.

We begin our treatment of state to output maps by deriving a heuristic characterization of such maps, then abstracting a definition from this and finally proceeding to study them in more detail.

Consider a pair (C, A) with $\mathbb{F}^m$ as state space and $\mathbb{F}^p$ as output space. The observability map is the map $\mathscr{O}_{(C,A)} : \mathbb{F}^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ defined, for $\xi \in \mathbb{F}^m$, by $\mathscr{O}_{(C,A)}\xi = \sum_{i=1}^{\infty} \frac{CA^{i-1}\xi}{z^i} = C(zI - A)^{-1}\xi$. Let $j : \mathbb{F}^m \to \mathbb{F}[z]^m$ be the natural embedding given by $j\xi = \xi$. The state space $\mathbb{F}^m$ has an $\mathbb{F}[z]$-module structure induced by A. Clearly, with $f = H_{C(zI-A)^{-1}}$, we have $(f \circ j)\xi = \pi_- C(zI - A)^{-1}\xi = C(zI - A)^{-1}\xi = \mathscr{O}_{(C,A)}\xi$ or $f \circ j = \mathscr{O}_{(C,A)}$. Since the observability map $\mathscr{O}_{(C,A)}$ is an $\mathbb{F}[z]$-homomorphism, so is $f \circ j$. This homomorphism is injective if and only if the pair (C, A) is observable.

This leads to the following definition.

Definition 2.1. A restricted input/output map $f : \mathbb{F}[z]^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is called a **state to output map**, abbreviated as state/output map, if $\mathbb{F}^n$ can be given an $\mathbb{F}[z]$-module structure such that, with $j : \mathbb{F}^n \to \mathbb{F}[z]^n$ the natural embedding defined by $j\xi = \xi$, we have $(f \circ j) : \mathbb{F}^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is an $\mathbb{F}[z]$-homomorphism. We say that f is an observable state/output map if it is a state/output map and the map $(f \circ j) : \mathbb{F}^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is injective.

Observable state/output maps are easily characterized. Before giving the characterization, we introduce a notational convention. Given an $\mathbb{F}$-linear space $\mathscr{X}$ and an $\mathbb{F}$-linear map $C : X \to \mathbb{F}^p$, it extends in a natural way to an $\mathbb{F}[z]$-module homomorphism of $z^{-1}X[[z^{-1}]] \to z^{-1}\mathbb{F}[[z^{-1}]]^p$, which we still denote by C, and which is given by

$$C\sum_{i=1}^{\infty} x_i z^{-i} = \sum_{i=1}^{\infty}(Cx_i)z^{-i}.$$

Theorem 2.5. *Given a restricted input/output map $f : \mathbb{F}[z]^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ corresponding to the strictly proper transfer function G. Let $j : \mathbb{F}^n \to \mathbb{F}[z]^n$ be the natural embedding defined by $j\xi = \xi$. Then the following statements are equivalent*:

1. *f is an observable state/output map.*
2. *There exists an observable pair (C, A) for which $f \circ j = \mathscr{O}_{(C,A)}$ or equivalently, $f = H_{C(zI-A)^{-1}}$.*
3. *For f, the map $(f \circ j) : \mathbb{F}[z]^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is injective and satisfies*
$$\operatorname{Im}(f \circ j) = \operatorname{Im} f. \tag{24}$$
4. *For any left coprime factorization $G = D^{-1}H$, the columns of H are a basis of the polynomial model X_D.*
5. *Given any restricted input/output map $g : \mathbb{F}[z]^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ for which*
$$\operatorname{Im} g \subset \operatorname{Im} f, \tag{25}$$
there exists a unique $B \in \mathbb{F}^{m\times n}$ such that $g = fB$.
6. *For every strictly proper G' satisfying*
$$\operatorname{Im} H_{G'} \subset \operatorname{Im} H_G, \tag{26}$$
there exists a unique B such that
$$G'(z) = G(z)B. \tag{27}$$

Proof. (1) $\Rightarrow$ (2)

Assume f is an observable state/output map. Thus $\mathbb{F}^n$ can be given an $\mathbb{F}[z]$-module structure. We define a linear transformation A in $\mathbb{F}^n$ by $A\xi = z \cdot \xi$. For $\xi \in \mathbb{F}^n$, let $(f \circ j)\xi = \sum_{k=1}^{\infty} \frac{\xi_k}{z^k}$.

Since the ξ_k depend linearly on ξ, there exist linear operators $L_k : \mathbb{F}^n \to \mathbb{F}^n$ for which $\xi_k = L_k\xi$, i.e. we have

$$(f \circ j)\xi = \sum_{k=1}^{\infty} \frac{\xi_k}{z^k} = \sum_{k=1}^{\infty} \frac{L_k \xi}{z^k}.$$

Using the fact that $f \circ j$ is an $\mathbb{F}[z]$-homomorphism, we compute

$$\sum_{k=1}^{\infty} \frac{L_k A\xi}{z^k} = (f \circ j)A\xi = (f \circ j)(z \cdot \xi) = z \cdot (f \circ j)\xi = \sigma \sum_{k=1}^{\infty} \frac{L_k \xi}{z^k} = \sum_{k=1}^{\infty} \frac{L_{k+1}\xi}{z^k}.$$

Denoting $C = L_1$ and equating coefficients, we have $L_{k+1} = L_k A$. By induction, we have $L_k = CA^{k-1}$ for $k \geqslant 1$. Substituting back, we have $(f \circ j) = \mathcal{O}_{(C,A)}$ or equivalently, $G = C(zI - A)^{-1}$. The observability of the pair (C, A) follows from the assumption that f is an observable state/output map. Finally, using the fact that f is an $\mathbb{F}[z]$-homomorphism, we compute

$$\begin{aligned} f \sum_k z^k \xi_k &= \sum_k z^k \cdot f(\xi_k) = \pi_- z^k \sum_k (f \circ j)(\xi_k) \\ &= \pi_- z^k \sum_k \cdot C(zI - A)^{-1}\xi_k = \pi_- C(zI - A)^{-1} \sum_k z^k \xi_k \\ &= H_{C(zI-A)^{-1}} \sum_k z^k \xi_k. \end{aligned}$$

This shows that $f = H_{C(zI-A)^{-1}}$.

(2) $\Rightarrow$ (3)

As $\mathcal{O}_{(C,A)} = H_{C(zI-A)^{-1}} j$, clearly $\operatorname{Im} \mathcal{O}_{(C,A)} \subset \operatorname{Im} H_{C(zI-A)^{-1}}$. To show the inverse inclusion, let $h \in \operatorname{Im} H_{C(zI-A)^{-1}}$, i.e. there exists a polynomial $u = \sum_k z^k \xi_k \in \mathbb{F}[z]^n$ for which $h = \pi_- C(zI - A)^{-1} u$. We compute

$$\begin{aligned} h &= \pi_- C(zI - A)^{-1} \sum_k z^k \xi_k = \sum_k \pi_- z^k C(zI - A)^{-1} \xi_k = \sum_k C(zI - A)^{-1} A^k \xi_k \\ &= \mathcal{O}_{(C,A)} \sum_k A^k \xi_k \in \operatorname{Im} \mathcal{O}_{(C,A)}. \end{aligned}$$

(3) $\Rightarrow$ (1)

Assume $(f \circ j) : \mathbb{F}^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is injective and satisfies (24). Note that $f : \mathbb{F}[z]^n \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ is an $\mathbb{F}[z]$-homomorphism, i.e. satisfies, for $u \in \mathbb{F}[z]^n$, $\sigma f(u) = f(zu)$. Let $\{e_1, \ldots, e_n\}$ be a basis for $\mathbb{F}^n$. We compute

$$\sigma(f \circ j)e_i = \sigma f(e_i) = f(ze_i) \in \operatorname{Im} f = \operatorname{Im} f \circ j.$$

So there exists a $\xi_i \in \mathbb{F}[z]^n$ for which $\sigma(f \circ j)e_i = (f \circ j)\xi_i$. Define $A : \mathbb{F}^n \to \mathbb{F}^n$ by $Ae_i = \xi_i$ and extend by linearity. Thus, for all $\xi \in \mathbb{F}^n$, we have $\sigma(f \circ j)\xi = (f \circ j)A\xi$. By induction and linearity, we have

$$p(\sigma)(f \circ j)\xi = (f \circ j)p(A)\xi, \tag{28}$$

which shows that, for the $\mathbb{F}[z]$-module structure on $\mathbb{F}^n$ induced by A, $f \circ j$ is an $\mathbb{F}[z]$-homomorphism, that is $f \circ j$ is a state/output map. By our assumption that $f \circ j$ is injective it follows that $f \circ j$ is an observable $\mathbb{F}[z]$-homomorphism.

(2) $\Rightarrow$ (4)

Let $f = H_G$ with transfer function $G(z) = C(zI - A)^{-1}$ for some observable pair (C, A), i.e. the input map is the identity I which is both injective and surjective. Let $D^{-1}H$ be any left coprime factorization of G. We consider the shift realization $\Sigma(D^{-1}H)$, given by (4), in the state space X_D. By the assumed left coprimeness, $\Sigma(D^{-1}H)$ is both reachable and observable. Applying the state space isomorphism theorem to the two realizations, the input map $B_\Sigma\xi = H(z)\xi$ is both injective and surjective. Injectivity means that the columns of $H(z)$ are linearly independent, whereas surjectivity that they span X_D. Taken together, this means that the columns of $H(z)$ are a basis of X_D.

(4) $\Rightarrow$ (6)

Assume that $G = D^{-1}H$ is a left coprime factorization and that the columns of $H(z)$ form a basis for X_D. Let G' be a $p \times q$ strictly proper transfer function for which the inclusion (26) holds. So, for every $u \in \mathbb{F}[z]^q$, we have $\pi_-G'u \in X^D = \operatorname{Ker} D(\sigma)$. Thus $0 = \pi_-D\pi_-G'u = \pi_-DG'u$. This implies that $N = DG'$ is a polynomial matrix and $G' = D^{-1}N$. The strict properness of G' implies that every column of N belongs to X_D and hence is a linear combination of the columns of $H(z)$. Thus there exists a constant matrix B for which $N = H(z)B$ and so $G'(z) = D(z)^{-1}N(z) = D(z)^{-1}H(z)B = G(z)B$.

(6) $\Rightarrow$ (5)

Write $g = H_{G_1}$ and $f = H_G$. Assume the inclusion (25) which is equivalent to (26). Thus there exists a unique B for which $G_1(z) = G(z)B$. We compute

$$g = H_{G_1} = H_{GB} = H_G B = fB.$$

(5) $\Rightarrow$ (4)

Write $f = H_G$ with $G = D^{-1}H$ a left coprime factorization and $g = H_{G_1}$. Any G_1 having a, not necessarily left coprime, left matrix fraction representation $G_1 = D^{-1}S$ satisfies $\operatorname{Im} H_{G_1} \subset \operatorname{Im} H_G$. We now choose S to be a basis matrix for the polynomial model X_D. Thus, there exists a unique B for which $D^{-1}S = D^{-1}HB$ or $S = HB$. Since the columns of S span X_D, so do the columns of H. Since B is uniquely determined, the columns of H are linearly independent and so also H is a basis matrix.

(4) $\Rightarrow$ (2)

Assume $G = D^{-1}H$ is a left coprime factorization with the columns of $H(z)$ a basis for X_D. Let (A, B, C) be an arbitrary minimal realization of G. By the state space isomorphism theorem, the shift realization $\Sigma(D^{-1}H)$ is isomorphic to (A, B, C). The input map of the realization $\Sigma(D^{-1}H)$ is given by $B_\Sigma\xi = H(z)\xi$. Since H is a basis matrix, B_Σ is both injective and surjective, i.e. invertible. By isomorphism, so is B. Therefore, we can rewrite the realization as $G(z) = C(zI - A)^{-1}B = (CB)(zI - (B^{-1}AB))^{-1}$. Adjusting notation, the implication is proved. $\square$

We prove now the following result, stated without proof in Proposition 3.4 of Fuhrmann [10].

Proposition 2.1. *The following statements are equivalent*:

1. *The behavior $\mathscr{B}$ is finite dimensional.*
2. $\mathscr{B} = \operatorname{Ker} D(\sigma)$ *for some nonsingular polynomial matrix* $D(z)$.

3. *$\mathscr{B}$ is equal to the rational model X^D.*
4. *There exists an observable pair (C, A) for which*

$$\mathscr{B} = \{C(zI - A)^{-1}\xi \,|\, \xi \in \mathbb{F}^n\} = \operatorname{Im} \mathscr{O}_{(C,A)}. \tag{29}$$

Proof. (1) ⇒ (2)

Assume $\mathscr{B}$ is finite dimensional. Thus, in a minimal kernel representation $\mathscr{B} = \operatorname{Ker} D(\sigma)$, the polynomial matrix $D(z)$ is necessarily nonsingular, otherwise there are some free variables, contradicting finite dimensionality.

(2) ⇒ (3)

Let $\mathscr{B} = \operatorname{Ker} D(\sigma)$ with $D(z)$ nonsingular. Let $X^D = \operatorname{Im} \pi^D$ with the projection π^D defined in the preliminaries. Let $h \in X^D$, i.e. $h = \pi_- D^{-1}\pi_+ Dh$. This implies

$$D(\sigma)h = \pi_- Dh = \pi_- D\pi_- D^{-1}\pi_+ Dh = \pi_- DD^{-1}\pi_+ Dh = \pi_-\pi_+ Dh = 0,$$

i.e. $X^D \subset \operatorname{Ker} D(\sigma)$.

Conversely, assume $h \in \operatorname{Ker} D(\sigma)$, i.e. $\pi_- Dh = 0$. This implies $\pi^D h = \pi_- D^{-1}\pi_+ Dh = \pi_- D^{-1} Dh = \pi_- h = h$, i.e. $h \in X^D$ or $\operatorname{Ker} D(\sigma) \subset X^D$. The two inclusions imply the equality $\mathscr{B} = \operatorname{Ker} D(\sigma) = X^D$.

(3) ⇒ (4)

Let $H(z)$ be a basis matrix for X_D. Let (A, B, C) be a minimal realization of $D^{-1}H$. Applying the state space isomorphism theorem as in the proof of Theorem 2.5, we have the invertibility of B and hence, redefining the matrices C and A, the equality $D(z)^{-1}H(z) = C(zI - A)^{-1}$. This implies the representation (29).

(4) ⇒ (1)

Let $\mathscr{B} = \{C(zI - A)^{-1}\xi \,|\, \xi \in \mathbb{F}^n\}$. Then, using the coprime factorization $D(z)^{-1}H(z) = C(zI - A)^{-1}$, we have $\mathscr{B} = X^D$ and hence $\dim \mathscr{B} = \deg \det D < \infty$. □

3. Markovian systems and first order representations

Classically, realization theory deals with the passage from an input/output representation of a system to a first order representation. Obviously, this is closely linked to the concept of state. The behavioral approach rejects input/output representations as a legitimate starting point for the definition of a linear system. Thus we are left with the problem of introducing the concept of state without leaning heavily on input/output, or transfer function, thinking. Before discussing states, we introduce and analyze the more general concept of Markovianity. Basically, given a dynamical system with behavior $\mathscr{B}$, what we are interested in is how much does the past of a signal influence its future. For some systems the past completely determines the future. However, it might be the case that less than the whole past determines the future of a trajectory. Thus the past of a trajectory is condensed into a smaller set of values a trajectory attains and the future of another trajectory of the behavior can be connected to the past of the first so long as they agree on the condensed set of values. This, on the behavioral level, is best described in terms of concatenation of signals. As a first step, following Willems [27], we introduce concatenations.

Since we work in the discrete time setting, with the time set being $\mathbb{Z}_+$, we need to modify a little the original definition of concatenation.

Definition 3.1. Given a behavior $\mathscr{B}$ and two trajectories $w^{(1)}, w^{(2)} \in \mathscr{B}$, we define their concatenation at time T, denoted by $w = w^{(1)} \wedge_T w^{(2)}$ by

$$(w^{(1)} \wedge_T w^{(2)})_t = \begin{cases} w_t^{(1)} & t \leqslant T, \\ w_t^{(2)} & t > T. \end{cases} \tag{30}$$

Note that

$$w^{(1)} \wedge_T w^{(2)} = w^{(2)} + [(w^{(1)} - w^{(2)}) \wedge_T 0]. \tag{31}$$

Hence, given $w^{(1)}, w^{(2)} \in \mathscr{B}$, then $w^{(1)} \wedge_T w^{(2)} \in \mathscr{B}$ if and only if $(w^{(1)} - w^{(2)}) \wedge_T 0 \in \mathscr{B}$ or $0 \wedge_T (w^{(1)} - w^{(2)}) \in \mathscr{B}$.

In general, a behavior is not closed under concatenation. For concatenability, certain compatibility conditions need to be satisfied and the study of those conditions is very much related to the concept of Markovianity, introduced next.

Definition 3.2

1. Let $\mathbb{F}$ be a field and l a positive integer. Given a discrete time dynamical system $(\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$, we say that it is
   (a) **l-Markovian** if for all $T \geqslant l$ and $w \in \mathscr{B}$, $w_T = w_{T-1} = \cdots = w_{T-l+1} = 0$ implies $w \wedge_T 0 \in \mathscr{B}$.
   (b) **strongly l-Markovian** if for all $T \geqslant l$ and $w \in \mathscr{B}$, $w_T = w_{T-1} = \cdots = w_{T-l+1} = 0$ implies $w = 0$.
2. Given a continuous time, dynamical system $(\mathbb{R}, \mathbb{R}^m, \mathscr{B})$, we say it is
   (a) **l-Markovian** if $w, \bar{w} \in \mathscr{B}$ and $w^{(i)}(0) = \bar{w}^{(i)}(0)$ for $i = 0, \ldots, l-1$, implies $w \wedge \bar{w} \in \mathscr{B}$.
   (b) **strongly l-Markovian** if $w \in \mathscr{B}$ and $w^{(i)}(0) = 0$ for $i = 0, \ldots, l-1$, implies $w = 0$.

A system is called **autonomous** if for some $l > 0$ it is strongly l-Markovian. If $\mathscr{B}$ is 1-Markovian, we will say it is **Markovian**.

Note that in the continuous time case the conditions for l-Markovianity and strong l-Markovianity are given at time $t = 0$. This is due to the fact that the time axis is $\mathbb{R}$ and the behavior $\mathscr{B}$ is translation invariant. This no longer applies to discrete time systems where the time set is $\mathbb{Z}_+$.

In Willems [26], l-Markovian systems are called systems with l-finite memory or systems with memory span l.

Clearly, the state system $w_{k+1} = Aw_k$ is Markovian. Indeed, any solution is given by $w_k = A^k w_0$ and if $w_k = 0$ then $A^{k+j} w_k = 0$. However, this does not exhaust Markovian systems. Suppose a system is defined by

$$\Sigma := \begin{cases} x_{k+1} = Ax_k, \\ y_k = Cx_k, \end{cases}$$

then with the behavior variable $w_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix}$, we have a first order system with behavior given by $\operatorname{Ker}\begin{pmatrix} \sigma I - A & 0 \\ C & -I \end{pmatrix}$ which is clearly a Markovian system. Note that Σ will be strongly Markovian if and only if A is invertible.

Consider next the system defined by

$$\Sigma := \begin{cases} x_{k+1} = Ax_k, \\ w_k = Cx_k, \end{cases}$$

as a system with latent variable x, manifest variable w and behavior given by

$$\begin{pmatrix} \sigma I - A & 0 \\ C & -I \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

If $D^{-1}H$ is a left coprime factorization of $C(zI - A)^{-1}$, then by elimination the behavior is given by $\operatorname{Ker} D(\sigma) = X^D$. It will be l-Markovian only if the pair (C, A) is observable and l is greater or equal to the largest observability index of the pair. Thus, in general, this system will not be Markovian. The only exception is when all row indices of $D(z)$ are bounded by 1. This is equivalent to all observability indices of the pair (C, A) being bounded by 1 which, if the pair (C, A) is observable, is equivalent to C being nonsingular.

Assume now that (C, A) is an observable pair. For Σ to be strongly l-Markovian we need

$$CA^{k+j}x = 0, \quad \text{for } j = 0, \dots, l-1 \tag{32}$$

to imply $CA^j x = 0$ for all $j \geqslant 0$. Now if $l \geqslant \nu$ where ν is the largest observability index of the observable pair (C, A), then if (32) is satisfied then $A^k x = 0$. We can conclude from this that, if A is invertible then necessarily $x = 0$ which shows that Σ is indeed strongly l-Markovian. Note that the invertibility of A is equivalent to $D(0)$ being nonsingular.

It is obvious that if a behavior is (strongly) l-Markovian, then it is (strongly) k-Markovian for all $k > l$.

We have the following simple result.

Lemma 3.1. *Given a discrete time dynamical system* $\Sigma = (\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$, *with* $\mathscr{B} = \operatorname{Ker} R(\sigma)$. *Assume* $R(z) = \begin{pmatrix} R_1(z) & R_2(z) \end{pmatrix}$. *Then the following statements hold*:

1. Σ *is l-Markovian implies that* $\Sigma_1 = (\mathbb{Z}_+, \mathbb{F}^{m_1}, \mathscr{B}_1)$, *with* $\mathscr{B}_1 = \operatorname{Ker} R_1(\sigma)$, *is l-Markovian.*
2. Σ *is strongly l-Markovian implies that* $\Sigma_1 = (\mathbb{Z}_+, \mathbb{F}^{m_1}, \mathscr{B}_1)$, *with* $\mathscr{B}_1 = \operatorname{Ker} R_1(\sigma)$, *is strongly l-Markovian.*
3. *Assume* $\mathscr{B} = \operatorname{Ker} R(\sigma)$ *with* $R(z) \in \mathbb{F}[z]^{p\times m}$ *of full row rank. A necessary and sufficient condition for Σ to be autonomous is that $\mathscr{B}$ is finite dimensional, i.e. that R is square and nonsingular.*

Proof

1. Let $w \in \operatorname{Ker} R_1(\sigma)$ and $w_1 = \cdots = w_l = 0$. Clearly $\begin{pmatrix} w \\ 0 \end{pmatrix} \in \operatorname{Ker} R(\sigma)$ and $\begin{pmatrix} w_1 \\ 0 \end{pmatrix} = \cdots = \begin{pmatrix} w_l \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Since Σ is assumed to be l-Markovian, we have $\begin{pmatrix} w \\ 0 \end{pmatrix} \wedge_l \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} w \wedge_l 0 \\ 0 \end{pmatrix} \in \mathscr{B}$. This shows that $w \wedge_l 0 \in \mathscr{B}_1$, i.e. $\mathscr{B}_1$ is l-Markovian.
2. Let $w \in \operatorname{Ker} R_1(\sigma)$ and $w_t = w_{t+1} = \cdots = w_{t+l-1} = 0$. Then $\begin{pmatrix} w \\ 0 \end{pmatrix} \in \operatorname{Ker} R(\sigma)$ and $\begin{pmatrix} w \\ 0 \end{pmatrix}_j = \begin{pmatrix} w_j \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for $j = t, \dots, t+l-1$. Since $\operatorname{Ker} R(\sigma)$ is strongly l-Markovian, it follows that $\begin{pmatrix} w_j \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for $j \geqslant t+l$. This implies of course that $w_j = 0$ for $j \geqslant t+l$, i.e. $\operatorname{Ker} R_1(\sigma)$ is strongly l-Markovian.
3. We begin by proving necessity. If R has full row rank and is properly rectangular, i.e. we have $m > p$, then by reordering the manifest variables, we can assume without loss of generality that $R(z) = \begin{pmatrix} Q(z) & -P(z) \end{pmatrix}$ with $Q(z)$ nonsingular and $Q^{-1}P$ proper. Splitting the manifest variables appropriately as $w = \begin{pmatrix} y \\ u \end{pmatrix}$, then $u \in z^{-1}\mathbb{F}[[z^{-1}]]^{m-p}$ can be freely chosen. Thus Σ cannot be autonomous.

   To prove sufficiency, assume that $\operatorname{Ker} R(\sigma)$ is finite dimensional, i.e. that R is square and nonsingular. Let $r(z) = \det R(z)$ and $\rho = \deg \det R(z)$. We will show, see also Lemma 6.2 in

Fuhrmann [10], that $X^R \cap z^{-(\rho+1)}\mathbb{F}[[z^{-1}]]^m = \{0\}$, i.e. X^R is strongly ρ-Markovian. To show this, let w be any element in the above intersection. Write $w = z^{-\rho}w'$ with $w' \in z^{-1}F[[z^{-1}]]^m$. Since $X^R \subset X^{\rho I}$, we compute

$$w = \pi^r w = \pi_- r^{-1}\pi_+ r w = \pi^r w = \pi_- r^{-1}\pi_+ r z^{-\rho} w' = 0,$$

as $rz^{-\rho}$ is proper. □

Note that, allowing permutation of the manifest variables, the statement of the lemma remains true for the projection onto an arbitrary subset of the manifest variables. Also, we point out that $\mathscr{B}_1 = \left\{\begin{pmatrix} w \\ 0 \end{pmatrix} \middle| w \in \operatorname{Ker} R_1(\sigma)\right\}$ is a subbehavior of $\mathscr{B} = \operatorname{Ker} R(\sigma)$, hence given via a factorization of R. It is easily checked that such a factorization is given by $(R_1(z) \quad R_2(z)) = \begin{pmatrix} I & 0 \end{pmatrix}\begin{pmatrix} R_1(z) & R_2(z) \\ 0 & I \end{pmatrix}$ and that $\operatorname{Ker}\begin{pmatrix} R_1(\sigma) & R_2(\sigma) \\ 0 & I \end{pmatrix} = \mathscr{B}_1$.

It follows from the Definition 3.2 that, both in the discrete and the continuous time case, strong l-Markovianity implies l-Markovianity. For autonomous, continuous time dynamical systems, the two concepts of Markovianity coincide.

Proposition 3.1. *Given an autonomous, continuous time dynamical system* $(\mathbb{R}, \mathbb{R}^m, \mathscr{B})$, *then it is strongly l-Markovian if and only if it is l-Markovian.*

Proof. It suffices to show that in this case l-Markovianity implies strong l-Markovianity. So, assume the system to be autonomous and l-Markovian. By (31) we have that $w \in \mathscr{B}$ and $w^{(i)} = 0$ for $i = 0, \ldots, l-1$ implies that $w \wedge 0 \in \mathscr{B}$. However, the elements of $\mathscr{B}$, all being exponential polynomials, are analytic, so $w(t) = 0$ for $t > 0$ implies $w = 0$. □

The equivalence of the concepts of l-Markovianity and that of strong l-Markovianity proved in Proposition 3.1 for autonomous, continuous time dynamical systems is no longer true in the discrete time case. In the case of discrete time systems, l-Markovianity does not necessarily imply strong l-Markovianity. To see this, consider $d(z) = z^l$, then for every $w \in \operatorname{Ker} \sigma^l$ and $k < l$, we have $w \wedge_k 0 \in \mathscr{B}$ but $w_1 = \cdots = w_k = 0$ does not imply $w = 0$.

However, something of Proposition 3.1 can still be saved. A polynomial matrix $E(z) \in \mathbb{F}[z]^{p\times k}$ will be called **monomic** if all its nonzero invariant factors, i.e. all its nonzero entries in its Smith canonical form, $\epsilon_1, \ldots, \epsilon_p$, are monomials, i.e. $\epsilon_i(z) = z^{\nu_i}$, with ν_i nonnegative. Since the determinant of a square polynomial matrix is the product of its invariant factors, a square polynomial matrix E is monomic if and only if $\det E(z) = z^n$ with $n = \sum_{i=1}^k \nu_i$. Clearly, a nonsingular polynomial matrix $D(z)$ admits a right monomic factor if and only if its determinant $d(z)$ has a monomic factor z^ν, or equivalently if and only if $d(0) = 0$. The next proposition shows that the existence of a monomic factor for $D(z)$ makes the difference between l-Markovianity and strong l-Markovianity.

Proposition 3.2. *Given a finite dimensional, discrete time dynamical system* $\Sigma = (\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$ *with* $\mathscr{B} = X^D$, *which is l-Markovian. Then it is strongly l-Markovian if and only if* $D(z)$ *has no monomic factor.*

Proof. To prove that if $\mathscr{B}$ is strongly l-Markovian then necessarily $D(z)$ has no monomic factor, we argue by contradiction. We show that if $D(z)$ has a monomic factor then it cannot be strongly l-Markovian. Indeed, if $D(z)$ has a monomic factor then there exists a nonzero $w \in \mathscr{B}$ such that $\sigma w = 0$, i.e. $w = \frac{w_1}{z}$. Thus $w_j = 0$ for $j \geqslant 2$ while $w \neq 0$, i.e. the system is not strongly l-Markovian.

Conversely, assume the system Σ is l-Markovian and that D has no monomic factor or, equivalently, that $d(0) = \det D(0) \neq 0$. Assume $w \in \mathscr{B}$ and $w_{T-l+1} = \cdots = w_T = 0$. By l-Markovianity, $\bar{w} = w \wedge_T 0 \in \mathscr{B}$, Clearly, for some positive integer N, we have $\sigma^N \bar{w} = 0$. Thus, there exists an element $0 \neq w' \in \mathscr{B}$ for which $\sigma w' = 0$, i.e. 0 is an eigenvalue of σ^D. But λ is an eigenvalue of σ^D if and only Ker $D(\lambda) \neq 0$ which is equivalent to $d(\lambda) = 0$. By our assumption on D, w' and hence also $\bar{w}$, is identically zero. This shows that Σ is strongly l-Markovian. □

Note that the absence of a monomic factor in $D(z)$ is equivalent to the behavioral equality $\sigma\mathscr{B} = \mathscr{B}$ while in general we have only $\sigma\mathscr{B} \subset \mathscr{B}$.

In the next proposition we study the relation between a class of continuous time behaviors and related, discrete time, ones. Since the analysis of continuous time behaviors is not the principal theme of this paper, we do not aim at a more general result.

For an analytic function f of exponential growth, we have the Taylor expansion $f(t) = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!} t^i$ and its Laplace transform given by $F(s) = \mathscr{L}(f) = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{s^{i+1}} = \sum_{i=0}^{\infty} \frac{F_i}{s^{i+1}}$. We shall interpret σ as differentiation in the time domain and as the backward shift in the frequency domain.

Proposition 3.3. *Let $T(z) \in \mathbb{C}[z]^{p\times p}$ be nonsingular. Let* Ker $T(\sigma) = \left\{ f | T\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) f = 0 \right\}$ *and X^T the rational model associated with T. Let*

1. *The Laplace transform is an $\mathbb{C}[z]$-homomorphism.*
2. *$\overline{\mathscr{L}}$, the restriction of the Laplace transform to* Ker $T(\sigma)$, *is a bijective map of* Ker $T(\sigma)$ *onto X^T.*
3. *Let (C^T, A^T) be defined by (6). Define the pair $(\widehat{C}_T, \widehat{A}_T)$ by*

$$\begin{cases} \widehat{A}_T \phi = D\phi = \phi' & \phi \in \operatorname{Ker} T(D), \\ \widehat{C}_T \phi = \phi(0) & \phi \in \operatorname{Ker} T(D). \end{cases} \tag{33}$$

Then the following diagram is commutative:

$$\begin{tikzcd}
\operatorname{Ker} T(\sigma) \arrow[r, "\hat{A}_T"] \arrow[dd, "\overline{\mathcal{L}}"'] & \operatorname{Ker} T(\sigma) \arrow[dd, "\overline{\mathcal{L}}"'] \arrow[dr, "\hat{C}_T"] & \\
 & & \mathbb{R}^p \\
X^T \arrow[r, "A^T"] & X^T \arrow[ur, "C^T"'] &
\end{tikzcd}$$

Proof

1. Follows from
$$\mathscr{L}(\sigma f) = \mathscr{L}(f') = s\mathscr{L}(f) - f(0) = \pi_- sF(s) = \sigma\mathscr{L}(f). \tag{34}$$
2. From (34), we obtain by induction
$$\mathscr{L}(\phi^{(j)}) = s^j \mathscr{L}(\phi) - s^{j-1}\phi(0) - \cdots - \phi^{(j-1)}(0) = s^j \mathscr{L}(\phi) - \sum_{\nu=0}^{j-1} s^{j-\nu-1}\phi^{(\nu)}(0).$$

For $T(z) = \sum_{j=0}^{r} T_j z^j$ and $\phi \in \operatorname{Ker} T(D)$, we have $T(D)\phi = 0$ and hence

$$\mathscr{L} \sum_{j=0}^{r} T_j \phi^{(j)} = 0 = \sum_{j=0}^{r} T_j \left[s^j \mathscr{L}(\phi) - \sum_{\nu=0}^{j-1} s^{j-\nu-1} \phi^{(\nu)}(0) \right],$$

or, using also a change of summation order, we get

$$\begin{aligned} T(s)\mathscr{L}(\phi) &= \sum_{j=0}^{r} \sum_{\nu=0}^{j-1} T_j s^{j-\nu-1} \phi^{(\nu)}(0) \\ &= \sum_{j=0}^{r-1} \sum_{\nu=j+1}^{r} T_j s^{j-\nu-1} \phi^{(\nu)}(0). \end{aligned}$$

From this we conclude that $\mathscr{L}$ maps $\operatorname{Ker} T(D)$ into X^T. It is easy to check, see in this connection Hinrichsen and Prätzel-Wolters [16], that the restricted map $\overline{\mathscr{L}} : \operatorname{Ker} T(\sigma) \to X^T$ is a bijection.

3. This is a simple verification. □

We proceed now to the characterization of l-Markovian systems. This is an extension of Proposition 3.1 in Rapisarda and Willems [21].

Recall first that, given $f = (f_1, \ldots, f_n) \in \mathbb{F}[z]^n$, with $\deg f = \max \deg f_i = \gamma$, we denote by $[f]_h$ the coefficient of z^γ in the expansion of f as a vector polynomial and γ the degree of f. Given a polynomial matrix $P \in \mathbb{F}[z]^{k \times n}$ with rows p_i, we define the **order** of P to be the sum of the row degrees. We also define $[P]_{hr} = ([f_1]_h, \ldots, [f_k]_h) \in \mathbb{F}^{k \times n}$ and call it the **highest row degree coefficient matrix**. A polynomial matrix P is called **row proper** if $[P]_{hr}$ has full row rank. Clearly, $[P]_{hr}$ has full row rank if and only if at least one of its $k \times k$ minors is nonzero. Thus, an $k \times n$ polynomial matrix P is row proper if and only if the maximum of the degrees of all $k \times k$ minors is equal to the order of P. It is well known, see Wolovich [29], that a full row rank $k \times n$ polynomial matrix P can be reduced to row proper form by elementary row operations, or equivalently by left multiplication by a unimodular polynomial matrix. The row degrees of a row reduced equivalent form are called the **row indices** of P and they are uniquely determined. For a reduction to the row proper, Kronecker–Hermite, canonical form, see Fuhrmann and Helmke [12].

Proposition 3.4

1. *Let $\Sigma := (\mathbb{Z}_+, \mathbb{F}^p, \mathscr{B})$ be an autonomous, discrete time system with behavior $\mathscr{B} = X^D$ and $D(z) \in \mathbb{F}[z]^{p \times p}$ nonsingular. Then Σ is l-Markovian if and only if $l \geqslant \nu$ with ν the largest row index of $D(z)$.*
2. *Given a system $\Sigma = (\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$ in the kernel representation $\mathscr{B} = \operatorname{Ker} R(\sigma)$ with $R \in \mathbb{F}[z]^{p \times m}$ of full row rank. Then the system is l-Markovian if and only if the row indices ν_i of R satisfy $l \geqslant \nu_1 \geqslant \cdots \geqslant \nu_p \geqslant 0$.*
3. *Let $\Sigma := (\mathbb{R}, \mathbb{R}^p, \mathscr{B})$ be an autonomous, continuous time system with behavior $\mathscr{B} = X^D$ and $D(z) \in \mathbb{R}[z]^{p \times p}$ nonsingular. Then Σ is l-Markovian if and only if $l \geqslant \nu$ with ν the largest row index of $D(z)$.*

Proof

1. Let $w \in X^D$, with $w = \sum_{k=1}^{\infty} \frac{w_k}{z^k}$. By Proposition 2.1, there exists an observable pair (C, A) for which $X^D = \{C(zI - A)^{-1}\xi | \xi \in \mathbb{F}^n\}$ and hence, for some constant vector ξ, we have $w_k = CA^{k-1}\xi$. Now the row indices of $D(z)$ are equal to the observability indices of the pair (C, A). In particular, ν is the largest observability index, i.e. the smallest integer for which, for an arbitrary vector ξ, $CA^j\xi = 0$ for $j = 0, \ldots, \nu - 1$ implies $\xi = 0$. Now if $l \geqslant \nu$ then, by the observability of (C, A), $CA^k\xi = \cdots = CA^{k+l-1}\xi = 0$ implies $A^k\xi = 0$ and hence $CA^{k+j}\xi = 0$ for all $j \geqslant 0$. This shows that Σ is l-Markovian.
   On the other hand, if $l < \nu$, then there exists a vector ξ for which $C\xi = \cdots = CA^{l-1}\xi = 0$ but $CA^l\xi \neq 0$. This shows that Σ is not l-Markovian.
2. The if part is immediate.
   To prove the converse, assume without loss of generality that $R(z)$ has full row rank and is row proper. By rearranging the behavioral variables, we can assume without loss of generality that $R(z) = \begin{pmatrix} Q(z) & -P(z) \end{pmatrix}$ with $Q(z)$ nonsingular and $Q^{-1}P$ proper. Clearly, the row indices of $Q(z)$ and $\begin{pmatrix} Q(z) & -P(z) \end{pmatrix}$ are the same. We split the behavioral variables accordingly as $w = \binom{y}{u}$. By Lemma 3.1, $X^Q = \operatorname{Ker} Q(\sigma)$ is also strongly l-Markovian. We may assume without loss of generality that $Q(z) = \operatorname{diag}(z^{\nu_1}, \ldots, z^{\nu_p})$ with $\nu_1 \geqslant \cdots \geqslant \nu_p$. Since we have the direct sum decomposition
   $$X^Q = X^{z^{\nu_1}} \oplus \cdots \oplus X^{z^{\nu_p}}, \tag{35}$$
   it suffices to show that if $\mathscr{B}$ is strongly l-Markovian, then $\nu_i \leqslant l$. In view of the direct sum representation (35), it suffices to show that X^{z^ν} is strongly l-Markovian implies $\nu \leqslant l$. To see this, note that $X^{z^\nu} = \{\sum_{i=1}^{\nu} \frac{\xi_i}{z^i} | \xi_i \in \mathbb{F}\}$. Clearly, $\xi_i = 0$ for $i = 1, \ldots, l$ implies $\xi_i = 0$ for $i = 1, \ldots, \nu$ if and only if $l \geqslant \nu$.
3. Using Proposition 3.3, the continuous time case is easily reducible to the discrete time case. Indeed, let $\phi \in \operatorname{Ker} T(D)$ and assume $l \geqslant \nu$ and that $\phi^{(j)}(0) = 0$ for $j = 0, \ldots, l-1$. This implies that $f = \overline{\mathscr{L}}\phi \in X^T$ satisfies $\sigma^j f = 0$ for $j = 0, \ldots, l-1$. Since X^T is l-Markovian, it follows that, necessarily, $f = 0$. As $\overline{\mathscr{L}}$ is an isomorphism, also $\phi = 0$, i.e. Σ is l-Markovian. This proves the sufficiency of the condition $l \geqslant \nu$ for l-Markovianity. Necessity is proved analogously by reduction to the discrete time case. □

An immediate, and important, corollary of Proposition 3.4 is the following characterization of Markovianity due to Rapisarda and Willems [21].

Proposition 3.5

1. *Given an autonomous system* $(\mathbb{Z}_+, \mathbb{F}^p, \mathscr{B})$ *in the kernel representation* $\mathscr{B} = \operatorname{Ker} R(\sigma)$. *Then the system is strongly Markovian if and only if* $\mathscr{B}$ *has a kernel representation of the form* $\mathscr{B} = \operatorname{Ker}(\sigma E + F)$ *for some constant square matrices* E, F *for which* $\det(zE + F) \neq 0$, *i.e.* $\mathscr{B}$ *is the kernel of a regular pencil.*
2. *Given a system* $(\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$ *in the kernel representation* $\mathscr{B} = \operatorname{Ker} R(\sigma)$. *Then the system is Markovian if and only if* $\mathscr{B}$ *has a kernel representation of the form* $\mathscr{B} = \operatorname{Ker}(\sigma E + F)$ *for some constant rectangular matrices* E, F *for which* $zE + F$ *has full row rank.*

In view of Proposition 3.3, the above result holds also for continuous time systems.

4. State systems and realization theory

The analysis of Markovianity was a prelude to the analysis of the concept of a state of a behavior. Recall that, other than in special cases like the physical modelling of a system, a linear system has been usually described in terms of the relations between inputs and outputs, be it via transfer functions or input/output relations, whereas in the behavioral setting by the trajectories of the manifest variables or via kernel representations, with or without the use of latent or auxiliary variables. In the Kalman approach to linear system theory, the state (or rather a state space description) is a construction based on input/output maps. The passage from external description to an internal model description, is referred to as realization theory. In order to emphasize the underlying unity of the various approaches to linear system theory, we shall keep the term realization also for the process of representing a behavior in terms of a first order system. To gain some intuition, we begin by looking at the case of autonomous behaviors.

4.1. Autonomous behaviors

Assume $D(z) \in \mathbb{F}[z]^{p\times p}$ is nonsingular. The corresponding autonomous behavior is given by $\mathscr{B} = \operatorname{Ker} D(\sigma) = X^D$. In order to obtain a first order representation for $\mathscr{B}$, we let $H(z)$ be an arbitrary basis matrix for the polynomial model X_D. In particular, $H(z)$ is a left prime polynomial matrix, i.e. has a polynomial right inverse. By Proposition 2.1, there exists a unique observable pair $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ such that

$$D(z)^{-1}H(z) = C(zI - A)^{-1}. \tag{36}$$

We claim now that the first order system, given by

$$\begin{pmatrix} I \\ 0 \end{pmatrix} w = \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} x \tag{37}$$

is a state space representation of the behavior $\mathscr{B} = X^D$ in the sense that $w \in \mathscr{B}$ if and only if there exists an $x \in z^{-1}\mathbb{F}[[z^{-1}]]^n$ such that (37) holds. In order to see this, we use Theorem 2.4 to eliminate the state variable from Eq. (37). Since, by the left primeness of $H(z)$, the factorizations in (36) are coprime, it follows that $\begin{pmatrix} D(z) & -H(z) \end{pmatrix}$ is a MLA annihilator of $\begin{pmatrix} C \\ zI - A \end{pmatrix}$. Hence, applying Theorem 2.4, the kernel representation of the behavior defined by (37) is given by

$$\operatorname{Ker} \begin{pmatrix} D(\sigma) & -H(\sigma) \end{pmatrix} \begin{pmatrix} I \\ 0 \end{pmatrix} = \operatorname{Ker} D(\sigma) = \mathscr{B}.$$

The observability of the pair (C, A) implies the existence of a polynomial left inverse of $\begin{pmatrix} C \\ zI - A \end{pmatrix}$. We now show that the observability of (C, A) implies that the state variable x is induced by the variable w, in the sense that there exists $X \in \mathbb{F}[z]^{n\times p}$ such that $x = X(\sigma)w$. Since the factorizations in (36) are coprime, we can embed the relation $H(z)(zI - A) = D(z)C$ in a doubly coprime factorization

$$\begin{aligned} \begin{pmatrix} D(z) & -H(z) \\ X(z) & Y(z) \end{pmatrix} \begin{pmatrix} \overline{Y}(z) & C \\ -\overline{X}(z) & zI - A \end{pmatrix} &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\ \begin{pmatrix} \overline{Y}(z) & C \\ -\overline{X}(z) & zI - A \end{pmatrix} \begin{pmatrix} D(z) & -H(z) \\ X(z) & Y(z) \end{pmatrix} &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \end{aligned} \tag{38}$$

Multiplying Eq. (37) by $\begin{pmatrix} D(\sigma) & -H(\sigma) \\ X(\sigma) & Y(\sigma) \end{pmatrix}$, and using the doubly coprime factorization (38), we obtain

$$\begin{pmatrix} D(\sigma) \\ X(\sigma) \end{pmatrix} w = \begin{pmatrix} D(\sigma) & -H(\sigma) \\ X(\sigma) & Y(\sigma) \end{pmatrix} \begin{pmatrix} I \\ 0 \end{pmatrix} w = \begin{pmatrix} D(\sigma) & -H(\sigma) \\ X(\sigma) & Y(\sigma) \end{pmatrix} \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} x = \begin{pmatrix} 0 \\ I \end{pmatrix} x, \tag{39}$$

i.e. we have

$$x = X(\sigma)w. \tag{40}$$

In fact, if $X(z)$ is defined via a doubly coprime factorization (38), then we can construct a behavior $\mathscr{B}_f$ by letting

$$\mathscr{B}_f = \begin{pmatrix} I \\ X(\sigma) \end{pmatrix} \mathscr{B}. \tag{41}$$

This of course can be equivalently written as the behavior $\mathscr{B}_f$ defined by

$$\begin{pmatrix} 0 \\ I \end{pmatrix} x = \begin{pmatrix} D(\sigma) \\ X(\sigma) \end{pmatrix} w, \tag{42}$$

with w considered here as a latent variable. Clearly, we have $\pi_w : \mathscr{B}_f \to \mathscr{B}$, defined by $\pi_w \begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = w$. Moreover, we compute for $w \in \mathscr{B}$, using the doubly coprime factorization (38) and the fact that $\mathscr{B} = \operatorname{Ker} D(\sigma)$,

$$\begin{aligned} \begin{pmatrix} C\sigma I - A \end{pmatrix} x &= \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} X(\sigma) w = \begin{pmatrix} CX(\sigma) \\ (\sigma I - A)X(\sigma) \end{pmatrix} w \\ &= \begin{pmatrix} I - \overline{Y}(\sigma)D(\sigma) \\ \overline{X}(\sigma)D(\sigma) \end{pmatrix} w = \begin{pmatrix} I \\ 0 \end{pmatrix} w, \end{aligned} \tag{43}$$

i.e. we have obtained a state representation of $\mathscr{B}$, given by (37). Following Rapisarda and Willems [21], we will call $X(z)$ a **state map**.

Note that from the doubly coprime factorization (38), we have in particular the intertwining relation

$$\overline{X}(z)D(z) = (zI - A)X(z). \tag{44}$$

By Theorem 2.2, it follows that the map $X(\sigma) : X^D \to X^{zI-A}$ is a B-isomorphism, i.e. that the following diagram commutes

$$\begin{tikzcd} X^D \arrow[r, "\sigma^D"] \arrow[d, "X(\sigma)"'] & X^D \arrow[d, "X(\sigma)"] \\ X^{zI-A} \arrow[r, "\sigma^{zI-A}"] & X^{zI-A} \end{tikzcd}$$

We note that $X^{zI-A} = \{\sum_{i=1}^{\infty} \frac{A^{i-1}\xi}{z^i} | \xi \in \mathbb{F}^n\}$ and $\sigma^{zI-A} \left(\sum_{i=1}^{\infty} \frac{A^{i-1}\xi}{z^i}\right) = A \sum_{i=1}^{\infty} \frac{A^{i-1}\xi}{z^i}$. This shows that the state map $X(\sigma)$ has linearized σ^D. The procedure outlined here is related to the Rosenbrock [22] definition of equivalence, dependent on matrix extensions and unimodular equivalence, or alternatively to the concept of linearization of polynomial matrices, see Gohberg et al. [14]. Given a nonsingular polynomial matrix $D(z)$ with $n = \deg \det D(z)$, we enlarge it to size $n \times n$ as to $\begin{pmatrix} D(z) & 0 \\ 0 & I \end{pmatrix}$ so that the extended matrix is unimodularly equivalent to a linear pencil $zI - A$.

We proceed next to adapt the classical, input/output based realization theory to the behavioral context. In the input/output context, a realization theory based on shift operators, see Fuhrmann [3], turned out to be an extremely efficient method at constructing realizations. The starting point of this method is a representation of a transfer function as left or right, not necessarily coprime, matrix fractions. This was extended, see Fuhrmann [4], to more general polynomial matrix descriptions, that is when a proper rational function is given in the form $VT^{-1}U + W$, or more generally, starting with a polynomial system matrix $\begin{pmatrix} T(z) & -U(z) \\ V(z) & W(z) \end{pmatrix}$. This type of representation is the starting point for the seminal work of Rosenbrock [22]. Since no coprimeness assumptions are made, then obviously one can assume some of the polynomial matrices U, V, W to be zero. Thus, in this formalism, we can study autonomous systems (U, V, W all zero), systems with no outputs (V zero), or systems with no inputs (U zero). In these cases there is no transfer function and the analysis of these systems did not fit easily in the standard realization theory. It is natural therefore to expect that the realization theory, put forward in Fuhrmann [3,4], could be adapted to the behavioral setting. This is indeed the case and we present here an extremely economical way of achieving this.

We extend this by formalizing the notions of state representation and state map. Our strategy is to develop the behavioral realization theory in analogy with the input/output approach to linear systems where the concept of state is directly linked to first order representations, i.e. to realization theory. Realization theory in that framework is developed in several stages which are:

1. Definition of state representations.
2. Existence of realizations.
3. Construction of realizations.
4. Definition and characterization of minimality.
5. Isomorphism theory for realizations.

Realizations are defined in terms of first order equations. The existence of realizations is proved via the factorization of the input/output map through a finitely generated torsion module that is a natural candidate for a state space based realization. On this level, the existence of a realization of an input/output map is characterized in terms of rationality, a result going back to Kronecker.

The construction of realizations depends on the representation of the transfer function. In terms of Markov parameters, there are algorithms for the construction, e.g. Kalman et al. [18]. More efficient methods are based on matrix fraction representations, see Fuhrmann [3,4].

Minimality of realizations is defined in terms of the dimension of the state space and characterized in terms of the properties of reachability and observability. Finally, for minimal realizations, we have the state space isomorphism theorem.

We intend to use this general structure as a road map for developing realization theory in the behavioral context.

4.2. *Definition of state representations*

State representations were introduced in Willems [25,26] and Rapisarda and Willems [21] directly in the behavioral setting. However, in order to gain some intuition, we consider the discrete time, state system with two trajectories satisfying, for $i = 1, 2$,

$$\begin{aligned} x_{j+1}^{(i)} &= Ax_j^{(i)} + Bu_j^{(i)}, \\ y_j^{(i)} &= Cx_j^{(i)} + Du_j^{(i)}. \end{aligned} \tag{45}$$

It is easy to check that if $\begin{pmatrix} y^{(i)} \\ u^{(i)} \\ x^{(i)} \end{pmatrix} \in \mathscr{B}_f$ and $x_T^{(1)} = x_T^{(2)}$, then $\begin{pmatrix} y^{(1)} \\ u^{(1)} \\ x^{(1)} \end{pmatrix} \wedge_T \begin{pmatrix} y^{(2)} \\ u^{(2)} \\ x^{(2)} \end{pmatrix} \in \mathscr{B}_f$. This leads us to the following.

Definition 4.1

1. A behavior $\mathscr{B}$ given in a latent variable system $(\mathbb{Z}_+, \mathbb{F}^m, \mathbb{F}^d, \mathscr{B}_f)$ is called a **state system** if $\begin{pmatrix} w^{(i)} \\ x^{(i)} \end{pmatrix} \in \mathscr{B}_f$ and $x_T^{(1)} = x_T^{(2)}$ implies $\begin{pmatrix} w^{(1)} \\ x^{(1)} \end{pmatrix} \wedge_T \begin{pmatrix} w^{(2)} \\ x^{(2)} \end{pmatrix} \in \mathscr{B}_f$.
2. A system $(\mathbb{Z}_+, \mathbb{F}^m, \mathbb{F}^d, \mathscr{B}_f)$ is a **state representation** of $\mathscr{B}$ if
   (a) It is a state system.
   (b) The projection $\pi_w : \mathscr{B}_f \to \mathscr{B}$, defined by $\pi_w \begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = w$

   is a behavior homomorphism of $\mathscr{B}_f$ onto $\mathscr{B}$. We say that d is the **dimension** of the state representation. A state representation of $\mathscr{B}$ is a **minimal state representation** if its dimension d is minimal among all state representations of $\mathscr{B}$.
3. A polynomial matrix $X(z) \in \mathbb{F}[z]^{d\times m}$ defines a **state map** $X(\sigma) : \mathscr{B} \to z^{-1}\mathbb{F}[[z^{-1}]]^d$ if
$$\mathscr{B}_f = \begin{pmatrix} I \\ X(\sigma) \end{pmatrix} \mathscr{B} \tag{46}$$
   is a state representation of $\mathscr{B}$. $X(\sigma)$ is a **minimal state map** if $\mathscr{B}_f$ defined by (46) is a minimal state representation of $\mathscr{B}$. The **dimension of a state map** is defined as the dimension of the corresponding realization.

The basic idea is that the representation $\mathscr{B}_f$ should be large enough so that it is a state representation, but small enough so that minimality is preserved. A full discussion of minimality will be given in Section 4.5.

The following proposition translates the previous, behavioral based, definition into algebraic terms:

Proposition 4.1. *Given the behavior* $\mathscr{B} = \operatorname{Ker} R(\sigma)$, *with* $R \in \mathbb{F}[z]^{p\times m}$ *of full row rank.*

1. *For arbitrary polynomial matrices* $R \in \mathbb{F}[z]^{p\times m}$ *and* $X \in \mathbb{F}[z]^{q\times m}$, *we have the identities*
$$\begin{pmatrix} I \\ 0 \end{pmatrix} R(z) = \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix} \begin{pmatrix} I \\ X(z) \end{pmatrix} \tag{47}$$
   and
$$R(z) \begin{pmatrix} I & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}. \tag{48}$$
2. *The identities* (47) *and* (48) *have the following doubly unimodular embedding*
$$\begin{aligned} &\begin{pmatrix} I & 0 & -R(z) \\ 0 & 0 & I \\ 0 & -I & X(z) \end{pmatrix} \begin{pmatrix} I & R(z) & 0 \\ 0 & X(z) & -I \\ 0 & I & 0 \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \\ &\begin{pmatrix} I & R(z) & 0 \\ 0 & X(z) & -I \\ 0 & I & 0 \end{pmatrix} \begin{pmatrix} I & 0 & -R(z) \\ 0 & 0 & I \\ 0 & -I & X(z) \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}. \end{aligned} \tag{49}$$

3. *The behavior* $\mathscr{B}_f = \begin{pmatrix} I \\ X(\sigma) \end{pmatrix}\mathscr{B} = \left\{ \begin{pmatrix} w \\ x \end{pmatrix} | w \in \mathscr{B} \& x = X(\sigma)w \right\}$ *has the kernel representation*

$$\mathscr{B}_f = \operatorname{Ker} \begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}. \tag{50}$$

Eliminating the latent variable x from the representation

$$\begin{pmatrix} R(\sigma) \\ X(\sigma) \end{pmatrix} w = \begin{pmatrix} 0 \\ I \end{pmatrix} x \tag{51}$$

gives the behavior $\mathscr{B} = \operatorname{Ker} R(\sigma)$.

4. *The map* $\begin{pmatrix} I \\ X(\sigma) \end{pmatrix} : X^R \to X^{\begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}}$ *is a B-isomorphism and its inverse is given by the map* $\begin{pmatrix} I & 0 \end{pmatrix} : X^{\begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}} \to X^R$.
5. *The polynomial matrix* $X \in \mathbb{F}[z]^{q\times m}$ *is a state map for* $\mathscr{B} = X^R$ *if and only if the embedding* $\operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}$ *is a state representation of* $\mathscr{B}$.

Proof

1. Checking (47) and (48) is a trivial computation.
2. Again, it is trivial to check.
3. Note that $\begin{pmatrix} w \\ x \end{pmatrix} \in \operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}$ if and only if $R(\sigma)w = 0$ and $x = X(\sigma)w$, i.e. $\mathscr{B}_f = \begin{pmatrix} I \\ X(\sigma) \end{pmatrix}\mathscr{B}$. Note that a MLA of $\begin{pmatrix} 0 \\ I \end{pmatrix}$ is $\begin{pmatrix} I & 0 \end{pmatrix}$. By the elimination theorem, we have that the manifest behavior of $\mathscr{B}_f$ is $\mathscr{B}$.
4. Follows from Theorem 2.3, using (49).
5. Follows from Definition 4.1. □

Note that if $X(z)$ induces a state map, and S is a nonsingular constant matrix, then also $SX(z)$ is a state map.

Proposition 4.1 shows that there are many embeddings of a behavior in latent variable systems which do not necessarily lead to state representations. To obtain characterizations of state representations and state maps, we need a clearer characterization of state systems and this is done next.

4.3. Characterization of realizations

In order to obtain a characterization of state systems, we use the characterization of Markovianity, given in Proposition 3.5. The following theorem was stated, without proof, in Rapisarda and Willems [21].

Theorem 4.1. *Let* $\mathscr{B}$ *have a latent variable representation* $(\mathbb{Z}_+, \mathbb{F}^m, \mathbb{F}^d, \mathscr{B}_f)$, *with latent variable* x. *Then this is a state representation of* $\mathscr{B}$ *if and only if there exist constant matrices* E, F, G *such that* $\mathscr{B}_f$ *has the kernel representation*

$$(\sigma E + F)x + Gw = 0. \tag{52}$$

Proof. If $\mathscr{B}$ has the first order representation (52), then clearly it is a state representation.

To prove the converse, assume $(\mathbb{Z}_+, \mathbb{F}^m, \mathbb{F}^d, \mathscr{B}_f)$ defines a state representation of $\mathscr{B}$. Since a state system is clearly Markovian, then, by Proposition 3.5, $\mathscr{B}_f$ has, for some constant matrices E, F, K, G, a kernel representation of the form $\mathscr{B}_f = \operatorname{Ker}\begin{pmatrix}\sigma K + G & \sigma E + F\end{pmatrix}$, i.e. we have

$$(\sigma K + G)w + (\sigma E + F)x = 0. \tag{53}$$

We want to show that there is an equivalent representation of the form (52).

We claim that, for $T \geqslant 1$, $\binom{0}{0} \wedge_T \binom{x}{w} \in \mathscr{B}_f$ if and only if

$$Ex(T) + Kw(T) = 0. \tag{54}$$

Indeed, concatenability at time T with the zero trajectory means, noting that $w(T-1) = 0$ and $x(T-1) = 0$ as they come from the zero trajectory, that

$$0 = Ex(T) + Fx(T-1) + Kw(T) + Gw(T-1) = Ex(T) + Kw(T),$$

and hence Eq. (54) holds. We show next that there exists a constant matrix $L \in \mathbb{F}^{q\times d}$ for which

$$Kw = Lx. \tag{55}$$

To this end, observe first that since x is a state variable, it follows that $x(T) = 0$ implies that $\binom{0}{0} \wedge_T \binom{x}{w} \in \mathscr{B}_f$, and hence, using (54), that $Kw(T) = 0$. We then define, for $T \geqslant 1$,

$$\mathscr{L}_T = \left\{ \begin{pmatrix} x(T) \\ w(T) \end{pmatrix} \middle| \begin{pmatrix} x \\ w \end{pmatrix} \in \mathscr{B}_f \right\}. \tag{56}$$

Now $\binom{x}{w} \in \mathscr{B}_f$ implies $\sigma\binom{x}{w} = \binom{\sigma x}{\sigma w} \in \mathscr{B}_f$, hence $\mathscr{L}_T \supset \mathscr{L}_{T+1}$, and, by induction, we have

$$\mathscr{L} = \mathscr{L}_1 \supset \mathscr{L}_2 \supset \cdots \tag{57}$$

Obviously, $\mathscr{L}$ is a finite dimensional vector space. We define two linear maps $\Phi : \mathscr{L} \to \mathbb{F}^d$ and $\Psi : \mathscr{L} \to \mathbb{F}^q$ by

$$\begin{aligned} \Phi\begin{pmatrix} x(1) \\ w(1) \end{pmatrix} &= x(1) \\ \Psi\begin{pmatrix} x(1) \\ w(1) \end{pmatrix} &= Kw(1). \end{aligned} \tag{58}$$

The implication $x(1) = 0 \Rightarrow Kw(1) = 0$ is equivalent to the kernel inclusion $\operatorname{Ker}\Phi \subset \operatorname{Ker}\Psi$. Thus, by a standard result in linear algebra, there exists a $q \times d$ matrix L for which $L\Phi = \Psi$, i.e.

$$Lx(1) = Kw(1). \tag{59}$$

Since $\mathscr{L}_T \subset \mathscr{L}$, for all T, it follows that $Lx(T) = Kw(T)$, so L extends to a map $Lx = Kw$ for all $\binom{x}{w} \in \mathscr{B}_f$. Substituting this relation in (53), we obtain $(E + L)\sigma x + Fx + Gw = 0$. Modifying the definition of E, the representation (52) follows. □

State representations and state maps were introduced, see Rapisarda and Willems [21], in the behavioral setting. However, for a full analysis, we need an algebraic characterization of state maps and the minimality of such maps. The next theorem is of central importance inasmuch as it provides such a characterization. Specifically, the characterization is given in terms of solvability of the polynomial equation (60). The construction of solutions to this equation will be given in Theorem 4.3.

Theorem 4.2. *Given a system* $\Sigma = (\mathbb{Z}_+, \mathbb{F}^m, \mathscr{B})$ *with behavior* $\mathscr{B} = \operatorname{Ker} R(\sigma)$ *and* $R(z) \in \mathbb{F}[z]^{p\times m}$ *of full row rank. Then*

1. $X(z) \in \mathbb{F}[z]^{d\times m}$ *is a state map for* Σ *if and only if there exist matrices* $E, F \in \mathbb{F}^{q\times d}$, $G \in \mathbb{F}^{q\times m}$ *and* $J(z) \in \mathbb{F}[z]^{q\times p}$ *right prime such that*
   (a) *The following equation is satisfied.*
   $$(zE + F)X(z) + G = J(z)R(z). \tag{60}$$
   (b) *Eliminating the latent variable x from the equation*
   $$(\sigma E + F)x + Gw = 0 \tag{61}$$
   yields the manifest behavior $\mathscr{B} = \operatorname{Ker} R(\sigma)$.
2. *The linear independence of the rows of a state inducing* $X(z)$ *is a necessary condition for the state map* $X(\sigma)$ *to be minimal.*
3. *If* $X(z)$ *is a minimal state map, then Eq.* (60) *can be rewritten in the form*
   $$\begin{pmatrix} R_\infty & -C \\ B & -(zI - A) \end{pmatrix} \begin{pmatrix} I \\ X(z) \end{pmatrix} = \begin{pmatrix} \overline{Y}_2 \\ -\overline{Y}_1 \end{pmatrix} R(z). \tag{62}$$
   Moreover, we have
   $$\begin{pmatrix} R_\infty \\ B \end{pmatrix} w = \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} x \tag{63}$$
   is a state representation of $\operatorname{Ker} R(\sigma)$.
4. *The observability of the pair* (C, A) *is a necessary condition for* $X(z)$ *to define a minimal state map.*

Proof

1. Assume there exist matrices $E, F \in \mathbb{F}^{q\times d}$, $G \in \mathbb{F}^{q\times m}$ and $J(z) \in \mathbb{F}[z]^{q\times p}$ right prime such that Eq. (60) holds. For an arbitrary $w \in \mathscr{B}$, we have $R(\sigma)w = 0$ and hence it follows that $(\sigma E + F)X(\sigma)w + Gw = 0$. This shows that $\begin{pmatrix} I \\ X(\sigma) \end{pmatrix}\mathscr{B} \subset \mathscr{B}_f = \left\{\begin{pmatrix} w \\ \xi \end{pmatrix} \mid (\sigma E + F)\xi + Gw = 0\right\}$. Thus, by Assumption 1b, we have that, with $\xi := X(\sigma)w$, the behavior $\mathscr{B}_f = \begin{pmatrix} I \\ X(\sigma) \end{pmatrix}\mathscr{B}$ is a state representation of $\mathscr{B}$.
   Conversely, assume $X(z)$ defines a state map, then, by Theorem 4.1, there exist matrices $E, F \in \mathbb{F}^{q\times d}$, $G \in \mathbb{F}^{q\times m}$ such that, for all $w \in \mathscr{B}$, $(\sigma E + F)X(\sigma)w + Gw = 0$. This implies the inclusion $\operatorname{Ker}[(\sigma E + F)X(\sigma) + G] \supset \operatorname{Ker} R(\sigma)$ which translates into a factorization (60). For $X(z)$ to define a state map, we need to have the equality $\operatorname{Ker}[(\sigma E + F)X(\sigma) + G] = \operatorname{Ker} R(\sigma)$. We show next that we have the equality if and only if in (60) the polynomial matrix $J(z)$ is right prime. To see this, assume $J(z)$ is right prime with $J^\sharp(z)$ an arbitrary polynomial left inverse. Then Eq. (60) implies $J^\sharp(z)((zE + F)X(z) + G) = R$, i.e. we have the inclusion $\operatorname{Ker}[(\sigma E + F)X(\sigma) + G] \subset \operatorname{Ker} R(\sigma)$. Since the opposite inclusion holds always, we have equality.
   Conversely, assume that $\operatorname{Ker}[(\sigma E + F)X(\sigma) + G] \subset \operatorname{Ker} R(\sigma)$, then there exists a polynomial matrix $L(z)$ for which $L(z)((zE + F)X(z) + G) = R(z)$. Using (60), we get $L(z)J(z)$

$R(z) = R(z)$. Since we assume that $R(z)$ has full row rank, it follows that $L(z)J(z) = I$, i.e. $J(z)$ is right prime.
Condition 4.2 follows from the fact that $x = X(\sigma)w$ is a state variable for $\mathscr{B}$, the external behavior described by (61).

2. Since $X(z)$ defines a state map, Eq. (60) is solvable. If the rows of $X(z)$ are not linearly independent, choose a nonsingular constant $d \times d$ matrix S so that $SX(z) = \begin{pmatrix} X_1(z) \\ 0 \end{pmatrix}$. Now

$$\begin{aligned}(zE+F)X(z)+G &= (zE+F)S^{-1}SX(z)+G = (zE+F)S^{-1}\begin{pmatrix} X_1(z) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} zE_1+F_1 & zE_2+F_2 \end{pmatrix}\begin{pmatrix} X_1(z) \\ 0 \end{pmatrix}+G \\ &= (zE_1+F_1)X_1(z)+G = J(z)R(z).\end{aligned}$$

Thus we have constructed a state map of smaller dimension.

3. Assume $X(z)$ is a minimal state map for $\mathscr{B} = \operatorname{Ker} R(\sigma)$. Thus Eq. (61) is a minimal first order representation of $\mathscr{B}$. We show that, necessarily, E has full column rank. If this is not the case, there exists a nonsingular, constant matrix T for which

$$\begin{aligned} ET &= \begin{pmatrix} E_1 & 0 \end{pmatrix} \\ FT &= \begin{pmatrix} F_1 & F_2 \end{pmatrix} \end{aligned}$$

with E_1 of full column rank. Letting $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = T^{-1}\xi$, the latent variable representation can be rewritten as $(\sigma E_1 + F_1)\xi_1 + F_2\xi_2 + Gw = 0$. Now N_2, the MLA of F_2 can be taken to be constant. Using a simple variant of Theorem 2.4, it follows that $(\sigma N_2 E_1 + N_2 F_1)\xi_1 + N_2 G w = 0$ is a state representation of $\mathscr{B}$ of lower dimension, contradicting minimality.
Since, by minimality, E has full column rank, there exist appropriately sized, nonsingular, constant matrices S, T for which we can write $SET = \begin{pmatrix} 0 & 0 \\ 0 & -I \end{pmatrix}$. Modifying the definitions of $X(z), G, J(z)$, the representation (62) follows.
To obtain (63), we use Eq. (62). For $w \in \operatorname{Ker} R(\sigma)$, we have

$$\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A) \end{pmatrix}\begin{pmatrix} I \\ X(\sigma) \end{pmatrix} w = \begin{pmatrix} \overline{Y}_2(\sigma) \\ -\overline{Y}_1(\sigma) \end{pmatrix} R(\sigma) w = 0,$$

which, with $x = X(\sigma)w$, implies (63).

4. If (C, A) is not observable, we have a nontrivial block representation

$$(C, A) = \left(\begin{pmatrix} C_1 & 0 \end{pmatrix}, \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}\right)$$

with (C_1, A_{11}) observable. Substituting in (62), we can rewrite it as

$$\begin{pmatrix} R_\infty & -C & 0 \\ B_1 & -(zI - A_{11}) & 0 \\ B_2 & A_{21} & -(zI - A_{22}) \end{pmatrix}\begin{pmatrix} I \\ X_1(z) \\ X_2(z) \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_{11}(z) \\ -\overline{Y}_{12}(z) \end{pmatrix} R(z).$$

Noting that $\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}$ is a MLA of $\begin{pmatrix} 0 \\ 0 \\ (zI - A_{22}) \end{pmatrix}$, we can apply Theorem 2.4 to the partial elimination of the X_2 variable to conclude that

$$\begin{pmatrix} R_\infty & -C \\ B_1 & -(zI - A_{11}) \end{pmatrix}\begin{pmatrix} I \\ X_1(z) \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_{11}(z) \end{pmatrix} R(z),$$

holds and that $\begin{pmatrix} R_\infty & -C \\ B_1 & -(\sigma I - A_{11}) \end{pmatrix} \begin{pmatrix} w \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a first order representation of $\mathscr{B} = \operatorname{Ker} R(\sigma)$. Therefore X_1 is a state map of dimension $d_1 < d$. $\square$

We wish to add that Kuijper [19] contains a comprehensive discussion of minimality of first order representations of behaviors. In particular Parts 3 and 4 of Theorem 4.2 are implied by the results in that reference.

It should be noted that Eq. (60) by itself does not guarantee that the corresponding $X(z)$ induces a state map. To see this consider the following simple example. Let $R(z) = z^2 + 3z + 1$. $X(z) = \begin{pmatrix} z+3 \\ 1 \end{pmatrix}$ is a state map for R. Indeed, with

$$R_\infty = 1, \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$
$$B = \begin{pmatrix} -1 \\ -3 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Eq. (60) becomes

$$\begin{pmatrix} 1 & 0 & -1 \\ -1 & -z & 0 \\ -3 & 1 & -z \end{pmatrix} \begin{pmatrix} 1 \\ z+3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} (z^2 + 3z + 1).$$

The equation

$$\begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix} w + \begin{pmatrix} 0 & -1 \\ -\sigma & 0 \\ 1 & -\sigma \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

is a latent variable representation of the behavior $\operatorname{Ker}(\sigma^2 + 3\sigma + 1)$ as can be checked by elimination. We still have

$$\begin{pmatrix} -1 \\ -3 \end{pmatrix} + \begin{pmatrix} -z & 0 \\ 1 & -z \end{pmatrix} \begin{pmatrix} z+3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} (z^2 + 3z + 1),$$

so Eq. (60) is satisfied with $J(z) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ right prime. However, the behavior given by the latent variable equation

$$\begin{pmatrix} -1 \\ -3 \end{pmatrix} w - \begin{pmatrix} \sigma & 0 \\ -1 & \sigma \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

represents, by the surjectivity of $\begin{pmatrix} \sigma & 0 \\ -1 & \sigma \end{pmatrix}$, the behavior $\mathscr{B} = z^{-1}\mathbb{F}[[z^{-1}]]$.

Note that Eq. (60) appears in Rapisarda and Willems [21], however the discussion there seems to be incomplete.

Corollary 4.1. *Given the behavior $\mathscr{B} = \operatorname{Ker} R(\sigma)$, with $R \in \mathbb{F}[z]^{p\times m}$ of full row rank. The polynomial matrix $X \in \mathbb{F}[z]^{q\times m}$ is a state map for X^R if and only if* $\operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}$ *has a* 1*st order representation.*

Proof. Follows from Theorem 4.1. $\square$

4.4. Existence and construction of realizations

While Proposition 4.1 shows that there are many ways of embedding a system in a latent variable systems, in general such an embedding does not provide a state representation. Our aim now is to find state representations.

So, we look now for conditions for an embedding of a system, whose behavior is given by the kernel representation

$$\mathscr{B} = \operatorname{Ker} R(\sigma), \tag{64}$$

in a first order latent variable system whose manifest behavior is equal to $\mathscr{B}$. Without loss of generality, we will assume that the kernel representation of $\mathscr{B}$ is minimal, i.e. that $R(z)$ is a $p \times m$ full row rank, row proper polynomial matrix with row indices $\nu_1 \geqslant \cdots \geqslant \nu_p$. We let $n = \sum_{i=1}^{p} \nu_i$. Assume now that $X(z)$ is a minimal state map. Thus we have a minimal first order representation $(\sigma E + F)x + Gw = 0$. By Theorem 4.2, this equation can be rewritten as $\begin{pmatrix} R_\infty & -C \\ B & -(zI - A) \end{pmatrix}\begin{pmatrix} I \\ X(z) \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix} R(z)$ and the latent variable behavior has two representations, namely $\operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}$ and $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A) \end{pmatrix}$. The assumption that $R(z)$ has full row rank implies that both representations are minimal. Necessarily, they are left unimodularly equivalent. Thus, there exists a unimodular polynomial matrix $\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}$ for which

$$\begin{pmatrix} U_{11}(z) & U_{12}(z) \\ U_{21}(z) & U_{22}(z) \end{pmatrix}\begin{pmatrix} R(z) & 0 \\ X & -I \end{pmatrix} = \begin{pmatrix} R_\infty & -C \\ B & -(zI - A) \end{pmatrix}. \tag{65}$$

Clearly, this implies that $\begin{pmatrix} U_{12}(z) \\ U_{22}(z) \end{pmatrix} = \begin{pmatrix} C \\ (zI - A) \end{pmatrix}$. We write $U_{11} = \overline{Y}_2, U_{21} = -\overline{Y}_1$. Letting $\begin{pmatrix} U_{11}(z) & U_{12}(z) \\ U_{21}(z) & U_{22}(z) \end{pmatrix}^{-1} = \begin{pmatrix} V_{11}(z) & V_{12}(z) \\ V_{21}(z) & V_{22}(z) \end{pmatrix}$, we conclude that $(V_{11}(z) \quad V_{12}(z))$ is a MLA of $\begin{pmatrix} C \\ (zI - A) \end{pmatrix}$. By Proposition 3.6 of Fuhrmann [10], the nonsingularity of $zI - A$ implies the nonsingularity of the polynomial matrix $V_{11}(z)$. Writing the unimodular matrix $V(z)$ as $V(z) = \begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix}$, it follows that $D(z)^{-1}H(z)$ is a left coprime factorization of $C(zI - A)^{-1}$. In particular, $H(z)$ is a basis matrix for X_D. From Eq. (65) we get, using the doubly coprime factorization

$$\begin{aligned} &\begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix}\begin{pmatrix} \overline{Y}_2(z) & C \\ -\overline{Y}_1(z) & zI - A \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\ &\begin{pmatrix} \overline{Y}_2(z) & C \\ -\overline{Y}_1(z) & zI - A \end{pmatrix}\begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \end{aligned} \tag{66}$$

the equality

$$\begin{pmatrix} R(z) & 0 \\ X & -I \end{pmatrix} = \begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix}\begin{pmatrix} R_\infty & -C \\ B & -(zI - A) \end{pmatrix}. \tag{67}$$

In turn this implies the equalities

$$\begin{aligned} R(z) &= D(z)R_\infty - H(z)B \\ X(z) &= Y_1(z)R_\infty + Y_2(z)B. \end{aligned} \tag{68}$$

From the first equation, using the strict properness of $D^{-1}H$, we get $D(z)^{-1}R(z) = R_\infty - D(z)^{-1}H(z)B$ and hence $R_\infty = \pi_+ D(z)^{-1}R(z)$. Taking this as a heuristic analysis, we are ready for the study of realizations and the construction of state maps for general behaviors. This is taken up next.

It is well known, see Willems [26,27], that by a rearrangement of the behavioral variables, we can assume that $R(z) = \begin{pmatrix} D(z) & -N(z) \end{pmatrix}$, with $D^{-1}N$ a proper rational function. However, in general, there is no uniqueness in such a procedure and there may be various orderings leading to different representations. Therefore, we find it preferable to take a more neutral approach. The following theorem provides a concrete, computational way to the construction of minimal state maps and addresses uniqueness issues. We believe that this theorem clarifies the relation between construction of state maps and realization theory.

Theorem 4.3. *Given a behavior $\mathscr{B}$ having the kernel representation (64), where $R(z)$ is a $p \times m$ full row rank, row proper polynomial matrix. Let $D(z)$ be any nonsingular $p \times p$, row proper polynomial matrix for which $D(z)^{-1}R(z)$ is proper and has a proper inverse. Let $H(z)$ be an arbitrary basis matrix for the polynomial model X_D. Let (C, A) be the unique observable pair for which*

$$D(z)^{-1}H(z) = C(zI - A)^{-1}. \tag{69}$$

holds. Let the intertwining relation $H(z)(zI - A) = D(z)C$ be embedded in a doubly coprime factorization

$$\begin{aligned} \begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix} \begin{pmatrix} \overline{Y}_2(z) & C \\ -\overline{Y}_1(z) & zI - A \end{pmatrix} &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\ \begin{pmatrix} \overline{Y}_2(z) & C \\ -\overline{Y}_1(z) & zI - A \end{pmatrix} \begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix} &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \end{aligned} \tag{70}$$

Then

1. *With R_∞, R_1 be defined by*
$$R_\infty = \pi_+ D^{-1}R, \tag{71}$$
and
$$R_1(z) = R(z) - D(z)R_\infty \tag{72}$$
respectively, there exists a unique, constant, $n \times m$ matrix B for which
$$R_1(z) = -H(z)B. \tag{73}$$
Thus we have the representation
$$R(z) = D(z)R_\infty - H(z)B. \tag{74}$$
2. *The behavior defined by the ARMA representation*
$$\begin{pmatrix} R_\infty \\ B \end{pmatrix} w = \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} x, \tag{75}$$
coincides with $\mathscr{B}$ and provides a first order representation for $\mathscr{B}$.
3. *Let $\begin{pmatrix} Y_1(z) & Y_2(z) \end{pmatrix}$ be as in the doubly coprime factorization (70), i.e. a solution to the Bezout equation*
$$Y_1(z)C + Y_2(z)(zI - A) = I. \tag{76}$$
Then, with $X(z)$ defined by
$$X(z) = Y_1(z)R_\infty + Y_2(z)B \tag{77}$$
$X(\sigma)$ is a state map, i.e.
$$\begin{pmatrix} R(\sigma) \\ X(\sigma) \end{pmatrix} w = \begin{pmatrix} 0 \\ I \end{pmatrix} x \tag{78}$$
is a state representation of $\mathscr{B}$.

4. *Let A, B, C, R_∞ be defined as above, and $X(z)$ defined by (77). Then*
   (a) *The matrices*

$$\begin{pmatrix} R_\infty & -C \\ B & -(zI-A) \end{pmatrix}, \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix} \tag{79}$$

   have full row rank and are left unimodularly equivalent. In fact, we have

$$\begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix}\begin{pmatrix} R_\infty & -C \\ B & -(zI-A) \end{pmatrix} = \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}. \tag{80}$$

   (b) *The behaviors defined by* $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I-A) \end{pmatrix}$ *and* $\operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I \end{pmatrix}$ *coincide.*
5. *With*

$$zE+F = \begin{pmatrix} -C \\ -(zI-A) \end{pmatrix}, \quad G = \begin{pmatrix} R_\infty \\ B \end{pmatrix}, \quad J(z) = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix}$$

   the equation

$$(zE+F)X(z)+G = J(z)R(z) \tag{81}$$

   holds, i.e. the following equation is satisfied:

$$\begin{pmatrix} -C \\ -(zI-A) \end{pmatrix} X(z) + \begin{pmatrix} R_\infty \\ B \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix} R(z). \tag{82}$$

Proof
1. By construction, $D(z)^{-1}R(z)$ is proper and $D(z)^{-1}R_1(z) = D(z)^{-1}R(z) - R_\infty$ strictly proper. Thus the columns of R_1 are in X_D and hence uniquely represented as linear combinations of the basis matrix columns.
2. Next, we will use the Theorem 2.4 to verify that a natural representation of the behavior is indeed a first order one. The coprimeness of the factorizations in (69), implies that the polynomial matrix $\begin{pmatrix} D(z) & -H(z) \end{pmatrix}$ is a MLA of $\begin{pmatrix} C \\ zI-A \end{pmatrix}$. Applying $\begin{pmatrix} D(\sigma) & -H(\sigma) \end{pmatrix}$ to both sides of Eq. (75) leads to

$$\begin{aligned} 0 &= \begin{pmatrix} D(\sigma) & -H(\sigma) \end{pmatrix}\begin{pmatrix} R_\infty \\ B \end{pmatrix} w = (D(\sigma)R_\infty - H(\sigma)B)w \\ &= (D(\sigma)R_\infty + R_1(\sigma))w = R(\sigma)w. \end{aligned}$$

   Using (72), this shows that indeed (75) is a first order representation of the behavior $\mathscr{B} = \operatorname{Ker} R(\sigma)$.
3. We show that

$$\begin{pmatrix} \overline{Y}_2(z) & C \\ -\overline{Y}_1(z) & (zI-A) \end{pmatrix}\begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix} = \begin{pmatrix} R_\infty & -C \\ B & -(zI-A) \end{pmatrix}. \tag{83}$$

   We compute first, using the doubly coprime factorization (70),

$$\begin{aligned} \overline{Y}_2(z)R(z) + CX(z) &= \overline{Y}_2(z)(D(z)R_\infty - H(z)B) + C(Y_1(z)R_\infty + Y_2(z)B) \\ &= (\overline{Y}_2(z)D(z) + CY_1(z))R_\infty + (CY_2(z) - \overline{Y}_2(z)H(z))B = R_\infty. \end{aligned}$$

The other relations are similarly computed.

4. Clearly, the polynomial matrix $\begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}$ has full row rank as R is assumed to have full row rank. Thus, it suffices to show that we have left unimodular equivalence. This follows, using the doubly coprime factorization (66), from the following computation.

$$\begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix} \begin{pmatrix} R_\infty & -C \\ B & -(zI-A) \end{pmatrix} = \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix}.$$

5. The second identity of the doubly coprime factorization (70) can be rewritten as

$$\begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix} (D(z) \quad -H(z)) + \begin{pmatrix} C \\ zI-A \end{pmatrix} (Y_1(z) \quad Y_2(z)) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

or, equivalently, as

$$\begin{pmatrix} -C \\ -(zI-A) \end{pmatrix} (Y_1(z) \quad Y_2(z)) + \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix} (D(z) \quad -H(z)).$$

Applying this equality to $\begin{pmatrix} R_\infty \\ B \end{pmatrix}$, and using (74) and (77), we obtain (82). □

4.5. Characterization of minimality

In preparation for the discussion of minimality of realizations, or state representations, for behaviors and the associated state maps as well as the analysis of minimality, we introduce, see Willems [27], the concept of McMillan degree for behaviors. We take as a model the standard definition of McMillan degree for proper rational functions. A proper rational function $G \in \mathbb{F}(z)^{p\times m}$ induces a Hankel operator $H_G : \mathbb{F}[z]^m \to z^{-1}\mathbb{F}[[z^{-1}]]^p$ defined by (2). The McMillan degree of G, $\delta(G)$, is defined as the smallest dimension n of a state representation $G(z) = D + C(zI-A)^{-1}B$, with A, B, C, D constant matrices of sizes $n \times n$, $n \times m$, $p \times n$, $p \times m$ respectively. Alternatively, see Fuhrmann [3], one can show that $\delta(G) = \dim \mathrm{Im}\, H_G$. Alternatively, if $G = Q^{-1}P$ is a left coprime factorization of G, then $\mathrm{Im}\, H_G = X^Q$ and so we have also $\delta(G) = \dim X^Q = \deg\det Q = \sum_{i=1}^p \nu_i$, where $\nu_1 \geqslant \cdots \geqslant \nu_p \geqslant 0$ are the row indices of Q (and hence of $(Q \quad -P)$). In view of this, one expects a similar characterization of the McMillan degree of a behavior and, indeed, this is achievable.

Clearly, since minimal kernel representation differ at most by a left unimodular factor, the degree of a behavior is well defined. Note also that if ν_i, $i = 1, \ldots, p$ are the row indices of $R(z)$, i.e. the row degrees of a row proper form of $R(z)$, then $\delta(\mathscr{B}) = n = \sum_{i=1}^p \nu_i$, see Fuhrmann [10].

To introduce the McMillan degree for behaviors, note that a behavior $\mathscr{B}$ has a kernel representation of the form $\mathscr{B} = \mathrm{Ker}\, R(\sigma)$, with $R(z) \in \mathbb{F}[z]^{p\times m}$ taken to be row proper and of full row rank. In this case, R is uniquely defined up to a constant, nonsingular left factor. With the polynomial matrix R, we associate the **reverse Hankel operator** $\mathscr{H}_R : z^{-1}\mathbb{F}_r[[z^{-1}]]^p \to \mathbb{F}_r[z]^m$ defined by

$$\mathscr{H}_R h = \pi_+ hR. \tag{84}$$

Here $z^{-1}\mathbb{F}_r[[z^{-1}]]^p$ and $\mathbb{F}_r[z]^m$ are defined as before, except that we are using row, rather than column, vectors. Clearly, by acting on elements of the form ηz^{-i}, we define

$$\Xi_R = \operatorname{Im} \mathscr{H}_R. \tag{85}$$

Both spaces $z^{-1}\mathbb{F}_r[[z^{-1}]]^p$ and $\mathbb{F}_r[z]^m$ have natural $\mathbb{F}[z^{-1}]$-module structures. In $z^{-1}\mathbb{F}_r[[z^{-1}]]^p$ this is given by multiplication, whereas in $\mathbb{F}_r[z]^m$ by the action $p \cdot f = \pi_+ pf = p(\sigma_+)f$, for $f \in \mathbb{F}_r[z]^m$ and $p \in \mathbb{F}[z^{-1}]$. Here σ_+ is defined by (3).

Since Hankel operators satisfy a functional equation, so do reversed Hankel operators. In fact, we have for $h \in z^{-1}\mathbb{F}_r[[z^{-1}]]^p$,

$$\begin{aligned}\mathscr{H}_R z^{-1}h &= \pi_+ R(z)z^{-1}h = \pi_+ z^{-1}R(z)h \\ &= \pi_+ z^{-1}\pi_+ R(z)h = \sigma_+\mathscr{H}_R h,\end{aligned}$$

i.e. we have the functional equation

$$\mathscr{H}_R z^{-1}h = \sigma_+\mathscr{H}_R h. \tag{86}$$

In particular, the functional equation implies $\sigma_+ \operatorname{Im} \mathscr{H}_R \subset \operatorname{Im} \mathscr{H}_R$, i.e. $X_R = \operatorname{Im} \mathscr{H}_R$ is σ_+-invariant.

Definition 4.2. Let $\mathscr{B}$ be a behavior in $z^{-1}\mathbb{F}[[z^{-1}]]^m$ having the kernel representation $\mathscr{B} = \operatorname{Ker} R(\sigma)$, with $R(z) \in \mathbb{F}[z]^{p\times m}$ taken to be of full row rank. We define the **McMillan degree** $\delta(\mathscr{B})$ of the behavior, by

$$\delta(\mathscr{B}) = \dim \Xi_R, \tag{87}$$

where X_R is defined by (85).

Note that $\Xi_R \subset \mathbb{F}_r[z]^m$ while $X_R \subset \mathbb{F}[z]^p$. Although X_R and Ξ_R are different spaces, they have the same dimension.

Proposition 4.2. *Let $R(z)$ be a $p \times m$ full row rank, row proper polynomial matrix with row indices $\nu_1 \geqslant \cdots \geqslant \nu_p \geqslant 0$ and let $n = \sum_{i=1}^p \nu_i$. With X_R defined by $X_R = R \operatorname{Ker} R(\sigma) = RX^R$ and Ξ_R defined by (85), we have*

$$\delta(R) = \dim X_R = \dim \Xi_R = n. \tag{88}$$

Proof. It is easy to check that in both cases, i.e. for X_R as well as Ξ_R, the dimension is equal to the sum of the row indices. $\square$

It would be nice to have a conceptual proof of this result.

Theorem 4.4. *With the notation and assumptions of Theorem 4.3, we have*

1. *There exists a doubly coprime factorization (70) with $\overline{Y}_1, \overline{Y}_2$ constant matrices.*
2. *$X(z) \in \mathbb{F}[z]^{d\times m}$ is a minimal state map if and only if the rows of $X(z)$ are a basis for the row space obtained from $R(z)$ by the shift down operation σ_+. Moreover, we have*

$$\delta(X) \geqslant \dim \Xi_R. \tag{89}$$

3. $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A)\end{pmatrix}$ *is a minimal realization of X^R if and only if the following conditions are satisfied:*

(a) *The pair* (C, A) *is observable.*
(b) *With* $D(z)^{-1}H(z)$ *a left coprime factorization of* $C(zI-A)^{-1}$, *we have the representation*

$$R(z) = D(z)R_\infty - H(z)B. \tag{90}$$

4. *The behavior defined by the ARMA representation*

$$\begin{pmatrix} R_\infty \\ B \end{pmatrix} w = \begin{pmatrix} C \\ \sigma I - A \end{pmatrix} x, \tag{91}$$

is a minimal first order representation for $\mathscr{B}$. *The state map* $X(z)$ *defined by* (77) *is a minimal state map.*

5. *The McMillan degree of the behavior* X^R *is equal to* $n = \dim \Xi_R = \sum_{i=1}^p \nu_i$.

Proof

1. Let $\begin{pmatrix} \overline{Y}'_2(z) \\ -\overline{Y}'_1(z) \end{pmatrix}$ be an arbitrary solution to the Bezout equation $D(z)\overline{Y}_2(z) + H(z)\overline{Y}_1(z) = I$. The general solution is given by

$$\begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix} = \begin{pmatrix} \overline{Y}'_2(z) \\ -\overline{Y}'_1(z) \end{pmatrix} + \begin{pmatrix} C \\ zI - A \end{pmatrix} Q(z)$$

with $Q(z)$ an arbitrary $n \times p$ polynomial matrix. Reducing $-\overline{Y}'_1$ modulo $zI - A$, we can write $-\overline{Y}'_1 = -\overline{Y}_1 - (zI-A)Q$, where $(zI-A)^{-1}\overline{Y}_1$ is strictly proper and $Q = \pi_+(zI - A)^{-1}\overline{Y}_1$. The strict properness of $(zI-A)^{-1}\overline{Y}_1$ means that $\overline{Y}_1$ is a constant matrix. We define now $\overline{Y}_2(z) = \overline{Y}'_2(z) + CQ(z)$. So we still have a solution to the Bezout equation with $\overline{Y}_1$ constant.
Next, we show that Y_1D^{-1} is strictly proper. Indeed, from the doubly coprime factorization (70), we have $\overline{Y}_1 D = (zI-A)Y_1$ or $Y_1D^{-1} = (zI-A)^{-1}\overline{Y}_1$. Since $\overline{Y}_1$ is constant, $(zI-A)^{-1}\overline{Y}_1$ is strictly proper, so also Y_1D^{-1} is strictly proper.
Finally, we consider the Bezout equation $\overline{Y}_2 D + CY_1 = I$ which implies $\overline{Y}_2 + CY_1D^{-1} = D^{-1}$. The row properness of D implies that D^{-1} is proper. Now $\overline{Y}_2$ is polynomial, D^{-1} is proper and Y_1D^{-1} strictly proper. From the previous equality we conclude that necessarily $\overline{Y}_2$ is a constant matrix.
2. Choose a constant solution $\overline{Y}_1, \overline{Y}_2$ of the Bezout equation $D\overline{Y}_2 - H\overline{Y}_1 = I$. Let $\begin{pmatrix} L_1 & L_2 \end{pmatrix}$ be a constant left inverse of $\begin{pmatrix} \overline{Y}_2 \\ -\overline{Y}_1 \end{pmatrix}$. Multiplying Eq. (82) by $\begin{pmatrix} L_1 & L_2 \end{pmatrix}$, we have $(L_1R_\infty + L_2B) - (L_1C + L_2(zI-A))X(z) = R(z)$. Applying the shift down map σ_+ to this equality, we get $\sigma_+R = -L_2X(z) - (L_1C - L_2A)\sigma_+X$. By induction, we conclude that, for all $j \geqslant 0$, $\sigma_+^j R$ is in the subspace spanned by $X, \sigma_+X, \ldots, \sigma_+^j X$. This shows that if $X(z)$ is a state map, then the subspace of $\mathbb{F}_r[z]^m$ spanned by the rows of $\sigma_+^j X$ has dimension $d > \dim \Xi_R = n$, i.e. we have inequality (89).
3. Assume the conditions (a) and (b) hold. Let $D(z)^{-1}H(z)$ be a left coprime factorization of $C(zI-A)^{-1}$ and (66) be an embedding of $D(z)C = H(z)(zI-A)$ in a doubly coprime factorization. We compute

$$\begin{aligned}
&\begin{pmatrix} D(z) & -H(z) \\ Y_1(z) & Y_2(z) \end{pmatrix} \begin{pmatrix} R_\infty & -C \\ B & -(zI-A) \end{pmatrix} \\
&\quad = \begin{pmatrix} D(z)R_\infty - H(z)B & 0 \\ Y_1(z)R_\infty + Y_2(z)B & -(Y_1(z)C + Y_2(zI-A)) \end{pmatrix} = \begin{pmatrix} R(z) & 0 \\ X(z) & -I \end{pmatrix},
\end{aligned}$$

i.e. $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A)\end{pmatrix} = \operatorname{Ker}\begin{pmatrix} R(\sigma) & 0 \\ X(\sigma) & -I\end{pmatrix}$. So $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A)\end{pmatrix}$ is a state representation of X^R. Now $\delta(X) = \dim \Xi_R = \sum_{i=1}^p \nu_i$, so this is indeed a minimal realization.
Conversely, assume that $\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A)\end{pmatrix}\begin{pmatrix} w \\ x\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix}$ is a minimal realization of X^R with $A \in \mathbb{F}^{n\times n}$. By Theorem 4.2, the pair (C, A) is necessarily observable. By Theorem 4.7 in Fuhrmann [3], we have

$$n = \deg\det(zI - A) = \deg\det D(z) = \dim X_D = \dim X_R = \dim \Xi_R.$$

Let $D(z)^{-1}H(z)$ be a left coprime factorization of $C(zI - A)^{-1}$. By elimination, i.e. applying Theorem 2.4, we have $X^R = \operatorname{Ker} R(\sigma) = \operatorname{Ker}(D(\sigma)R - H(\sigma)B)$. Both $R(z)$ and $(D(z)R - H(z)B)$ have full row rank, so they differ by at most a left unimodular factor. Since a left coprime factorization is determined only up to a common left unimodular factor, we may assume without loss of generality that (90) holds.

4. For the state map X defined by (77), we have $\delta(X) = n = \sum_{i=1}^p \nu_i$. On the other hand, from Part 2, it follows that for an arbitrary state map $X\prime$ we have $\delta(X') \geqslant \dim \Xi_R = n$. The minimality of X follows. □

Theorem 4.4 shows that the generator matrix A, even in the case of a minimal state representation of a behavior, is far from being uniquely determined. This result may seem counterintuitive inasmuch as in standard realization theory the spectral properties of the generating matrix A are completely determined by the singularities of the corresponding transfer function. To see that there is no contradiction, we review the realization of rational matrices. Let us assume that a rational matrix function G has a, not necessarily coprime, left matrix fraction representation $G = D^{-1}N$. Define $D_\infty = \pi_+ D^{-1}N$. Then $N = DD_\infty + N_1$, with $D^{-1}N_1$ strictly proper. Choosing a basis matrix $H(z)$ for the polynomial model X_D, there exists a unique observable pair (C, A) for which $D^{-1}H = C(zI - A)^{-1}$. Moreover, there exists a unique, constant matrix B for which $N_1(z) = H(z)B$. We compute

$$\begin{aligned} D_\infty + C(zI - A)^{-1}B &= D_\infty + D^{-1}H(z)B \\ &= D(z)^{-1}[D(z)D_\infty + H(z)B] = D(z)^{-1}N(z) = G(z), \end{aligned}$$

i.e. we have a realization of G. Thus a behavioral first order representation of the behavior $\operatorname{Ker}\begin{pmatrix} D(\sigma) & -N(\sigma)\end{pmatrix}$ is given by $\begin{pmatrix} I & -D_\infty & C \\ 0 & B & \sigma I - A\end{pmatrix}\begin{pmatrix} y \\ u \\ x\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix}$. To see this, we use elimination theory to eliminate the state variable x. We do this by applying $\begin{pmatrix} D(\sigma) & -H(\sigma)\end{pmatrix}$ to get

$$0 = \begin{pmatrix} D(\sigma) & -(D(\sigma)D_\infty + H(\sigma)B)\end{pmatrix}\begin{pmatrix} y \\ u\end{pmatrix} = \begin{pmatrix} D(\sigma) & -N(\sigma)\end{pmatrix}\begin{pmatrix} y \\ u\end{pmatrix}.$$

Note that $\begin{pmatrix} I & -D_\infty & C \\ 0 & B & \sigma I - A\end{pmatrix}\begin{pmatrix} y \\ u \\ x\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix}$ is the behavioral representation of the discrete time system

$$\begin{aligned} x_{j+1} &= Ax_j + Bu_j, \\ y_j &= Cx_j + D_\infty u_j, \end{aligned}$$

which can be written alternatively as

$$\begin{aligned} \sigma x &= Ax + Bu, \\ y &= Cx + D_\infty u. \end{aligned}$$

4.6. Isomorphism theory for realizations

A central result in the input/output theory of linear systems is Kalman's state space isomorphism theorem, see Kalman et al. [18], stating that any two minimal realizations of a given transfer function are isomorphic. It is expected that an analogous result holds in the behavioral context. This indeed turns out to be the case and the result is summed up in the following theorem.

Theorem 4.5. *With the notation of Theorem 4.3, we have*

1. *Two first order, minimal systems* $\operatorname{Ker}\begin{pmatrix} R_\infty & -C_i \\ B_i & -(\sigma I - A_i) \end{pmatrix}$, $i = 1, 2$, *represent the same behavior* $\mathscr{B} = \operatorname{Ker} R(\sigma)$, *if and only if there exist an output injection* $J \in \mathbb{F}^{n \times p}$ *and a nonsingular* $S \in \mathbb{F}^{n \times n}$ *such that*

$$\begin{pmatrix} R_\infty & -C_2 \\ B_2 & -(zI - A_2) \end{pmatrix} = \begin{pmatrix} R_\infty & -C_1 S \\ S^{-1}B + JR_\infty & -S^{-1}(zI - A_1 + JC_1)S \end{pmatrix} \tag{92}$$

an output injection map $J \in \mathbb{F}^{n \times p}$ *such that*

$$\operatorname{Ker}\begin{pmatrix} R_\infty & -C_2 \\ B_2 & -(\sigma I - A_2) \end{pmatrix} = \operatorname{Ker}\begin{pmatrix} R_\infty & -C_1 \\ B + JR_\infty & -(\sigma I - A_1 + JC_1) \end{pmatrix}.$$

Moreover, for the respective state maps defined by (77), *we have* $X_2(z) = X_1(z)$.

2. *Given two basis matrices* $H_1(z)$ *and* $H_2(z)$ *for* X_D. *Denote by* X_1, X_2 *the corresponding state maps. Then*
   (a) *There exists a constant nonsingular matrix S for which* $H_2(z) = H_1(z)S$.
   (b) *The doubly coprime factorization* (70) *transforms into*

$$\begin{pmatrix} D(z) & -H_1(z)S \\ S^{-1}Y_1(z) & S^{-1}Y_2(z)S \end{pmatrix}\begin{pmatrix} \overline{Y}_2(z) & CS \\ -S^{-1}\overline{Y}_1(z) & zI - S^{-1}AS \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{93}$$

   (c) *Let* $S \in \mathbb{F}^{n \times n}$ *be nonsingular. If* $\operatorname{Ker}\begin{pmatrix} R_\infty & C \\ B & \sigma I - A \end{pmatrix}$ *is a state representation for* $\mathscr{B}$, *then so is* $\operatorname{Ker}\begin{pmatrix} R_\infty & CS \\ S^{-1}B & \sigma I - S^{-1}AS \end{pmatrix}$.
   (d) *We have*

$$X_2(z) = S^{-1}X_1(z). \tag{94}$$

Proof

1. Let $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B & -(\sigma I - A) \end{pmatrix}$ be a state representation of the behavior $\mathscr{B}$. The behavior remains unchanged if the matrix $\begin{pmatrix} R_\infty & -C \\ B & -(zI - A) \end{pmatrix}$ is multiplied on the left by a unimodular polynomial matrix. Choosing the unimodular polynomial matrix $\begin{pmatrix} I & 0 \\ -J & I \end{pmatrix}$ we get the state representation $\operatorname{Ker}\begin{pmatrix} R_\infty & -C \\ B - JR_\infty & -(\sigma I - A - JC) \end{pmatrix}$.

Starting with the doubly coprime factorization (70), we get, inserting the factorization

$$\begin{pmatrix} I & 0 \\ J & I \end{pmatrix}\begin{pmatrix} I & 0 \\ -J & I \end{pmatrix}=\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

the doubly coprime factorization

$$\begin{pmatrix} D(z)-H(z)J & -H(z) \\ Y_1(z)+Y_2(z)J & Y_2(z) \end{pmatrix}\begin{pmatrix} \overline{Y}_2(z) & -C \\ \overline{Y}_1(z)-J\overline{Y}_2(z) & -(zI-A-JC) \end{pmatrix}=\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Computing the state map X_J, using (77), we have

$$\begin{aligned} X_J(z) &= (Y_1(z)+Y_2(z)J)R_\infty + Y_2(z)(B-JR_\infty) \\ &= Y_1(z)R_\infty + Y_2(z)B = X(z), \end{aligned}$$

i.e. the state map remains invariant.

Conversely, assume $\mathrm{Ker}\begin{pmatrix} R_\infty & -C_1 \\ B_1 & -(\sigma I-A_1) \end{pmatrix}$ and $\mathrm{Ker}\begin{pmatrix} R_\infty & -C_2 \\ B_2 & -(\sigma I-A_2) \end{pmatrix}$ are two minimal state representations of the behavior $\mathrm{Ker}\, R(\sigma)$ with state maps X_1, X_2 respectively. We use Eq. (80) for both representations, i.e. we have

$$\begin{pmatrix} D_i(z) & -H(z) \\ Y_1^{(i)}(z) & Y_2^{(i)}(z) \end{pmatrix}\begin{pmatrix} R_\infty & -C_i \\ B_i & -(zI-A_i) \end{pmatrix}=\begin{pmatrix} R(z) & 0 \\ X_i(z) & -I \end{pmatrix}, \quad i=1,2. \tag{95}$$

In particular, we have $D_1(z)R_\infty - H(z)B_1 = D_2(z)R_\infty - H(z)B_2$, or $(D_2(z)-D_1(z))R_\infty = H(z)(B_2-B_1)$. Since R_∞ has full row rank, it has a right inverse. Thus, there exists a matrix J such that

$$D_2(z) = D_1(z) + H(z)J. \tag{96}$$

From (95), we have the equalities $D_i(z)C_i = H(z)(zI-A_i)$. Subtracting, and using (96), we have $(D_1(z)-H(z)J)C_2 = H(z)(zI-A_2)$, or $D_1(z)C_2 = H(z)(zI-A_2+JC_2)$. Subtracting $D_1(z)C_1 = H(z)(zI-A_1)$ we obtain $D_1(z)(C_2-C_1) = H(z)(A_1-A_2+JC_2)$. Now $D_1^{-1}H$ is strictly proper. Thus we have

$$\begin{aligned} &C_2 = C_1 \\ &A_2 - JC_2 = A_1. \end{aligned} \tag{97}$$

Referring to (95) once more, we have $D_i(z)R_\infty - H(z)B_i = R(z)$. Subtracting, and using (96), we have $(D_1(z)-H(z)J)R_\infty - H(z)B_2 = D_1(z)R_\infty - H(z)B_1$, or $-H(z)JR_\infty = H(z)(B_2-B_1)$. As $H(z)$ is a basis matrix, this implies

$$B_2 = B_1 - JR_\infty. \tag{98}$$

It remains to verify the equality of the state maps. To this end we compare the two Bezout identities, using (97) and setting $C = C_1 = C_2$,

$$\begin{aligned} &Y_1^{(1)}C + Y_2^{(1)}(zI-A_2+JC) = I \\ &Y_1^{(2)}C + Y_2^{(2)}(zI-A_2) = I. \end{aligned}$$

Note that as a result of our assumption that all row indices ν_i of $R(z)$ are positive, the matrix C has full row rank. Thus we conclude from the previous equations that

$$\begin{aligned} &Y_1^{(2)} = Y_1^{(1)} + Y_2^{(1)}J \\ &Y_2^{(2)} = Y_2^{(1)}. \end{aligned} \tag{99}$$

Computing X_i by (77), we get

$$X_2 = Y_1^{(2)} R_\infty + Y_2^{(2)} B_2 = (Y_1^{(1)} + Y_2^{(1)} J) R_\infty + Y_2^{(1)} (B_1 - J R_\infty)$$
$$= Y_1^{(1)} R_\infty + Y_2^{(1)} B_1 = X_1.$$

2. (a) This is elementary linear algebra.
   (b) A result of a simple computation.
   (c) We use the elimination theorem once more.
   Note that $\begin{pmatrix} CS \\ zI - S^{-1}AS \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} C \\ zI - A \end{pmatrix} S$. Since S is invertible, an MLA of $\begin{pmatrix} CS \\ zI - S^{-1}AS \end{pmatrix}$ is obtained from a MLA of $\begin{pmatrix} C \\ zI - A \end{pmatrix}$ by the right factor $\begin{pmatrix} I & 0 \\ 0 & S \end{pmatrix}$. Now a MLA of $\begin{pmatrix} C \\ zI - A \end{pmatrix}$ is given by $\begin{pmatrix} D(z) & -H(z) \end{pmatrix}$. This implies that $\begin{pmatrix} D(z) & -H(z)S \end{pmatrix} = \begin{pmatrix} D(z) & -H(z) \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & S \end{pmatrix}$ is a MLA of $\begin{pmatrix} CS \\ zI - S^{-1}AS \end{pmatrix}$. We compute now, using the representation (74),

$$\begin{pmatrix} D(z) & -H(z)S \end{pmatrix} \begin{pmatrix} R_\infty \\ S^{-1}B \end{pmatrix} = D(z)R_\infty - H(z)B = R(z).$$

   This proves the statement.
   (d) We compute, using the doubly coprime factorization (70),

$$\begin{pmatrix} R(z) & 0 \\ X_2(z) & I \end{pmatrix} = \begin{pmatrix} D(z) & -H_1(z)S \\ S^{-1}Y_1(z) & S^{-1}Y_2(z)S \end{pmatrix} \begin{pmatrix} R_\infty & CS \\ S^{-1}B & zI - S^{-1}AS \end{pmatrix}$$
$$= \begin{pmatrix} D(z)R_\infty - H_1(z)B & (D(z)C - H_1(z)(zI - A))S \\ S^{-1}(Y_1(z)R_\infty + Y_2(z)B) & S^{-1}(Y_1(z)C + Y_2(z)(zI - A))S \end{pmatrix}$$
$$= \begin{pmatrix} R(z) & 0 \\ S^{-1}X_1(z) & I \end{pmatrix}. \qquad \square$$

4.7. A special basis

We saw in Theorem 4.3 that there is a freedom in computing a state map for a behavior Ker $R(\sigma)$. This freedom stems from the fact that, in the notation of that theorem, we can choose freely a nonsingular polynomial matrix D for which $D^{-1}R$ is proper as well as choose freely a basis for the space X_D arbitrarily. In this subsection we show that a natural choice is to take D in polynomial Brunovsky form and choose in X_D the standard basis discussed in Section 2.2. This leads to a simplified construction of a state map, see Rapisarda and Willems [21]. Assuming that the polynomial matrix $R(z)$ is in row proper form, the construction is essentially computation free. It is related to the construction of the control basis in Fuhrmann [5, 8] and to the realizations procedures given in Wolovich [29] and Schumacher and Rosenthal [23].

Proposition 4.3
1. *Let*

$$D(z) = \mathrm{diag}(z^{\nu_1}, \ldots, z^{\nu_p}) \tag{100}$$

and the standard basis for X_D *given, with* $h_i(z) = \begin{pmatrix} 1 & z & \cdot & \cdot & z^{\nu-1} \end{pmatrix}$, *by*

$$H(z) = \operatorname{diag}(h_1(z), \ldots, h_p(z)) = \begin{pmatrix} H^{(1)}(z) & \cdot & \cdot & \cdot & H^{(p)}(z) \end{pmatrix}$$
$$= \left(\begin{array}{ccccc|ccc|ccccc} 1 & z & \cdot & \cdot & z^{\nu_1-1} & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 1 & z & \cdot & \cdot & z^{\nu_p-1} \end{array}\right). \tag{101}$$

Let (C, A) *be the unique observable pair for which* (69) *holds. Let the* $1 \times \nu$ *and* $\nu \times \nu$ *matrices* L_ν *and* N_ν *by*

$$L_\nu = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix}$$
$$N_\nu = \begin{pmatrix} 0 & & & & & 0 \\ 1 & \cdot & & & & \cdot \\ & \cdot & \cdot & & & \cdot \\ & & \cdot & \cdot & & \cdot \\ & & & \cdot & \cdot & 0 \\ & & & & 1 & 0 \end{pmatrix}. \tag{102}$$

Let the $\nu \times 1$ *and* $\nu \times \nu$ *polynomial matrices* Z_ν *and* K_ν *be given by*

$$\begin{pmatrix} Z_\nu & | & K_\nu \end{pmatrix} = \left(\begin{array}{c|cccccc} z^{\nu-1} & 0 & -1 & -z & \cdot & \cdot & -z^{\nu-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -z \\ z & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{array}\right) \tag{103}$$

and the $1 \times \nu$ *and* $\nu \times \nu$ *polynomial matrices* $\overline{K}_\nu$ *and* $\overline{Z}_\nu$ *by*

$$\begin{pmatrix} \overline{K}_\nu \\ -\overline{Z}_\nu \end{pmatrix} = \left(\begin{array}{c} 0 \\ \hline -1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array}\right). \tag{104}$$

Then an associated doubly coprime factorization is given by (70), *where*

$$\begin{aligned} C &= \operatorname{diag}(L_{\nu_1}, \ldots, L_{\nu_p}) \\ zI - A &= \operatorname{diag}(zI - N_{\nu_1}, \ldots, zI - N_{\nu_p}) \\ Y_1(z) &= \operatorname{diag}(Z_{\nu_1}, \ldots, Z_{\nu_p}) \\ Y_2(z) &= \operatorname{diag}(K_{\nu_1}, \ldots, K_{\nu_p}) \\ \overline{Y}_1(z) &= \operatorname{diag}(\overline{Z}_{\nu_1}, \ldots, \overline{Z}_{\nu_p}) \\ \overline{Y}_2(z) &= \operatorname{diag}(\overline{K}_{\nu_1}, \ldots, \overline{K}_{\nu_p}). \end{aligned} \tag{105}$$

2. *Given a behavior $\mathscr{B}$ having the kernel representation (64), i.e. by $\mathscr{B} = \mathrm{Ker}\, R(\sigma)$, where $R(z)$ is a $p \times m$ full row rank, row proper polynomial matrix with row indices $\nu_1 \geqslant \cdots \geqslant \nu_p > 0$. We let $n = \sum_{i=1}^{p} \nu_i$. Let the rows of R be denoted by $R^{(i)}$, $i = 1, \ldots, p$. With σ_+ defined by (3), let $R_j^{(i)} = \sigma_+^j R^{(i)}$ for $j = 1, \ldots, \nu_i$. Then*

$$X(z) = \begin{pmatrix} R_1^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ R_{\nu_1}^{(1)} \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline R_1^{(p)} \\ \cdot \\ \cdot \\ \cdot \\ R_{\nu_p}^{(p)} \end{pmatrix} \tag{106}$$

is a minimal state map for $\mathscr{B}$.

Proof

1. We compute first the simple case where $d_\nu(z) = z^\nu$ and $h_\nu(z) = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & z^{\nu-1} \end{pmatrix}$. Applying Theorem 4.3, the rational matrix $d_\nu^{-1} h_\nu$ is realized by the pair (L_ν, N_ν). Indeed, it is easy to check that the following is a doubly coprime factorization.

$$\begin{pmatrix} d_\nu(z) & -h_\nu(z) \\ Z_\nu & K_\nu \end{pmatrix} \begin{pmatrix} \overline{K}_\nu & L_\nu \\ -\overline{Z}_\nu & zI - N_\nu \end{pmatrix}$$
$$= \left(\begin{array}{c|cccccc} z^\nu & -1 & -z & \cdot & \cdot & \cdot & -z^{\nu-1} \\ \hline z^{\nu-1} & 0 & -1 & -z & \cdot & \cdot & -z^{\nu-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -z \\ z & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{array}\right) \left(\begin{array}{c|cccccc} 0 & 0 & \cdot & \cdot & \cdot & \cdot & 1 \\ \hline -1 & z & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & -1 & z \end{array}\right)$$
$$= \left(\begin{array}{c|ccccc} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \hline 0 & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 & 1 \end{array}\right) = \begin{pmatrix} 1 & 0 \\ 0 & I \end{pmatrix}.$$

The general case follows now from the block diagonal structure of all matrices, given in Eqs. (100), (101) and (105).

2. With R_∞, R_1 and B defined as in (71), (72) and (73) respectively, we can write $H(z)B = \sum_{i=1}^{p} H^{(i)}(z) B_i$.

Computing the state map by (77), we obtain

$$
\begin{aligned}
X(z) &= Y_1(z)R_\infty + Y_2(z)B \\
&= \begin{pmatrix} Z_{\nu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & Z_{\nu_p} \end{pmatrix} \begin{pmatrix} R_\infty^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ R_\infty^{(p)} \end{pmatrix} + \begin{pmatrix} K_{\nu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & K_{\nu_p} \end{pmatrix} \begin{pmatrix} B_1 \\ \cdot \\ \cdot \\ \cdot \\ B_p \end{pmatrix} \\
&= \begin{pmatrix} Z_{\nu_1} R_\infty^{(1)} + K_{\nu_1}(z)B_1 \\ \cdot \\ \cdot \\ \cdot \\ Z_{\nu_p} R_\infty^{(p)} + K_{\nu_p}(z)B_p \end{pmatrix}
\end{aligned}
$$

which proves (106). $\square$

5. Examples

The first example treats the autonomous case. It also shows that we do not necessarily have to use the standard basis.

Example 1. Assume $d(z) = (z-\alpha)(z-\beta)$. $\mathscr{B} = \operatorname{Ker}(\sigma-\alpha)(\sigma-\beta) = \operatorname{span}\left\{\frac{1}{z-\alpha}, \frac{1}{z-\beta}\right\}$.

Choose a basis matrix $H(z) = \begin{pmatrix} z-\beta & z-\alpha \end{pmatrix}$ for X_d. The matrix representation of the shift realization of the stae/output map with respect to the chosen basis is given by $C = \begin{pmatrix} 1 & 1 \end{pmatrix}$, $A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$.

Embedding in a doubly coprime factorization, we have

$$
\begin{pmatrix} (z-\alpha)(z-\beta) & -(z-\beta) & -(z-\alpha) \\ \frac{1}{\alpha-\beta}(z-\beta) & -\frac{1}{\alpha-\beta} & -\frac{1}{\alpha-\beta} \\ -\frac{1}{\alpha-\beta}(z-\alpha) & \frac{1}{\alpha-\beta} & \frac{1}{\alpha-\beta} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ -\frac{1}{\alpha-\beta} & (z-\alpha) & 0 \\ \frac{1}{\alpha-\beta} & 0 & (z-\beta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

$$
\begin{pmatrix} 0 & 1 & 1 \\ -\frac{1}{\alpha-\beta} & (z-\alpha) & 0 \\ \frac{1}{\alpha-\beta} & 0 & (z-\beta) \end{pmatrix} \begin{pmatrix} (z-\alpha)(z-\beta) & -(z-\beta) & -(z-\alpha) \\ \frac{1}{\alpha-\beta}(z-\beta) & -\frac{1}{\alpha-\beta} & -\frac{1}{\alpha-\beta} \\ -\frac{1}{\alpha-\beta}(z-\alpha) & \frac{1}{\alpha-\beta} & \frac{1}{\alpha-\beta} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

By Eqs. (42) and (43), a state map is given by

$$
X(z) = \begin{pmatrix} \frac{1}{\alpha-\beta}(z-\beta) \\ -\frac{1}{\alpha-\beta}(z-\alpha) \end{pmatrix}.
$$

To see that

$$
\begin{pmatrix} (\sigma-\alpha)(\sigma-\beta) \\ \frac{1}{\alpha-\beta}(\sigma-\beta) \\ -\frac{1}{\alpha-\beta}(\sigma-\alpha) \end{pmatrix} w = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Xi_1 \\ \Xi_2 \end{pmatrix} \tag{107}
$$

is a state representation, we show that it is equivalent to a first order representation. This follows, using the doubly coprime factorization, that leads to

$$\begin{pmatrix}1\\0\\0\end{pmatrix} w = \begin{pmatrix}-(\sigma-\beta) & -(\sigma-\alpha)\\ -\frac{1}{\alpha-\beta} & -\frac{1}{\alpha-\beta}\\ \frac{1}{\alpha-\beta} & \frac{1}{\alpha-\beta}\end{pmatrix}\begin{pmatrix}\Xi_1\\ \Xi_2\end{pmatrix}.$$

Example 2. Consider $R(z) = \begin{pmatrix} z^2 + p_1 z + p_0 & -q_1 z - q_0\end{pmatrix}$. It is easily checked that, choosing $d(z) = z^2$, we have

$$R_\infty = \begin{pmatrix}1 & 0\end{pmatrix},$$
$$R_1(z) = \begin{pmatrix}p_1 z + p_0 & -q_1 z - q_0\end{pmatrix}.$$

We choose the standard basis in X_{z^2}, i.e. $H(z) = \begin{pmatrix}1 & z\end{pmatrix}$. Using (72), we have

$$\begin{pmatrix}p_1 z + p_0 & -q_1 z - q_0\end{pmatrix} = -\begin{pmatrix}1 & z\end{pmatrix}\begin{pmatrix}-p_0 & q_0\\ -p_1 & q_1\end{pmatrix}$$

i.e. $B = \begin{pmatrix}-p_0 & q_0\\ -p_1 & q_1\end{pmatrix}$. Using the matrix representation of the shift realization with respect to the chosen basis, we have

$$C = \begin{pmatrix}0 & 1\end{pmatrix},$$
$$A = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}.$$

By Theorem 4.3,

$$\begin{pmatrix}1 & 0\\ p_0 & -q_0\\ p_1 & -q_1\end{pmatrix}\begin{pmatrix}w_1\\ w_2\end{pmatrix} = \begin{pmatrix}0 & 1\\ \sigma & 0\\ -1 & \sigma\end{pmatrix}\begin{pmatrix}\Xi_1\\ \Xi_2\end{pmatrix} \tag{108}$$

is a state representation of the behavior $\operatorname{Ker}\begin{pmatrix}\sigma^2 + p_1\sigma + p_0 & q_1\sigma + q_0\end{pmatrix}$.
We can check this, invoking Theorem 2.4, by elimination of the latent variables. Note that $\begin{pmatrix}z^2 & -1 & -z\end{pmatrix}$ is a MLA of $\begin{pmatrix}0 & 1\\ z & 0\\ -1 & z\end{pmatrix}$, we get $\begin{pmatrix}z^2 & -1 & -z\end{pmatrix}\begin{pmatrix}1 & 0\\ p_0 & -q_0\\ p_1 & -q_1\end{pmatrix} = \begin{pmatrix}z^2 + p_1 z + p_0 & q_1 z + q_0\end{pmatrix}$, i.e. we get back our initial behavior. Since, obviously, $\begin{pmatrix}0\\ -1\\ 0\end{pmatrix}$ is a right inverse of $\begin{pmatrix}z^2 & -1 & -z\end{pmatrix}$, the Bezout Eq. (76) is solved by $\overline{Y}_1(z) = \begin{pmatrix}1\\0\end{pmatrix}$ and $\overline{Y}_2(z) = 0$. Inverting the matrix $\begin{pmatrix}\overline{Y}_2(z) & C\\ -\overline{Y}_1(z) & zI - A\end{pmatrix}$ yields $\begin{pmatrix}D(z) & -H(z)\\ Y_1(z) & Y_2(z)\end{pmatrix} = \begin{pmatrix}z^2 & -1 & -z\\ z & 0 & -1\\ 0 & -1 & z\end{pmatrix}$ and hence

$$Y_1(z) = \begin{pmatrix}z\\ 1\end{pmatrix}, \quad Y_2(z) = \begin{pmatrix}0 & -1\\ 0 & 0\end{pmatrix}.$$

Thus we compute the state map, using Eq. (77), to get

$$X(z) = \begin{pmatrix}z\\ 1\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}0 & -1\\ 0 & 0\end{pmatrix}\begin{pmatrix}-p_0 & q_0\\ -p_1 & q_1\end{pmatrix} = \begin{pmatrix}z + p_1 & -q_1\\ 1 & 0\end{pmatrix}.$$

Of course, this is in perfect agreement with Proposition 4.3, applying the shift and cut algorithm to $R(z) = \begin{pmatrix}z^2 + p_1 z + p_0 & -q_1 z - q_0\end{pmatrix}$.

Example 3. Consider $R(z) = \begin{pmatrix} z^2+p_1z+p_0 & q_1z+q_0 & r_2z^2+r_1z+r_0 \\ s_0 & z+t_0 & u_0 \end{pmatrix}$. This matrix is row proper with row indices 2, 1. Clearly $R_\infty = \begin{pmatrix} 1 & 0 & r_2 \\ 0 & 1 & 0 \end{pmatrix}$.

We choose therefore

$$D(z) = \begin{pmatrix} z^2 & 0 \\ 0 & z \end{pmatrix}, \quad H(z) = \begin{pmatrix} 1 & z & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using (72), we have

$$-H(z)B = \begin{pmatrix} p_1z+p_0 & q_1z+q_0 & r_1z+r_0 \\ s_0 & t_0 & u_0 \end{pmatrix},$$

and hence

$$B = \begin{pmatrix} -p_0 & -q_0 & -r_0 \\ -p_1 & -q_1 & -r_1 \\ -s_0 & -t_0 & -u_0 \end{pmatrix}.$$

Using the shift realization as in the previous example, we obtain

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and so a first order representation of the behavior $\mathscr{B} = \operatorname{Ker} R(\sigma)$ is given by

$$\begin{pmatrix} 1 & 0 & r_2 \\ 0 & 1 & 0 \\ -p_0 & -q_0 & -r_0 \\ -p_1 & -q_1 & -r_1 \\ -s_0 & -t_0 & -u_0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ z & 0 & 0 \\ -1 & z & 0 \\ 0 & 0 & z \end{pmatrix} \begin{pmatrix} \Xi_1 \\ \Xi_2 \\ \Xi_3 \end{pmatrix}.$$

We check this, using the elimination theorem.

Using the coprime factorizations $C(zI-A)^{-1} = D(z)^{-1}H(z)$, it follows that a MLA of the matrix $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ z & 0 & 0 \\ -1 & z & 0 \\ 0 & 0 & z \end{pmatrix}$ is $\begin{pmatrix} z^2 & 0 & -1 & -z & 0 \\ 0 & z & 0 & 0 & -1 \end{pmatrix}$. Computing

$$\begin{pmatrix} z^2 & 0 & -1 & -z & 0 \\ 0 & z & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & r_2 \\ 0 & 1 & 0 \\ -p_0 & -q_0 & -r_0 \\ -p_1 & -q_1 & -r_1 \\ -s_0 & -t_0 & -u_0 \end{pmatrix}$$
$$= \begin{pmatrix} z^2+p_1z+p_0 & q_1z+q_0 & r_2z^2+r_1z+r_0 \\ s_0 & z+t_0 & u_0 \end{pmatrix}$$

shows that, indeed, we have a state realization of $\mathscr{B}$.

Next, we solve the Bezout equation $D(z)\overline{Y}_2(z) + H(z)\overline{Y}_1(z) = I$. Obviously, we have

$$\begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} z^2 & 0 & -1 & -z & 0 \\ 0 & z & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -z^2 & 0 & 1 & z & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -z & 0 & 0 & 1 \end{pmatrix}.$$

Inverting the unimodular matrix leads to

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -z^2 & 0 & 1 & z & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -z & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ z^2 & 0 & 1 & -z & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & z & 0 & 0 & 1 \end{pmatrix}.$$

Now

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ z^2 & 0 & 1 & -z & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & z & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \overline{Y}_2(z) \\ -\overline{Y}_1(z) \end{pmatrix},$$

i.e.

$$\overline{Y}_1(z) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \overline{Y}_2(z) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The doubly coprime factorization becomes

$$\left(\begin{array}{cc|ccc} z^2 & 0 & -1 & -z & 0 \\ 0 & z & 0 & 0 & -1 \\ \hline z & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \hline -1 & 0 & z & 0 & 0 \\ 0 & 0 & -1 & z & 0 \\ 0 & -1 & 0 & 0 & z \end{array}\right) = \left(\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

which implies

$$Y_1(z) = \begin{pmatrix} z & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Y_2(z) = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Computing the state map, using Eq. (77), we get

$$X(z) = \begin{pmatrix} z & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & r_2 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -p_0 & -q_0 & -r_0 \\ -p_1 & -q_1 & -r_1 \\ -s_0 & -t_0 & -u_0 \end{pmatrix}$$
$$= \begin{pmatrix} z + p_1 & q_1 & r_2 z + r_1 \\ 1 & 0 & r_2 \\ 0 & 1 & 0 \end{pmatrix}$$

which is the matrix obtained from applying repeatedly the shift and cut operation to $R(z)$.

Acknowledgments

The authors would like to thank Jan Willems for some useful comments during the writing of this paper.

References

[1] E. Emre, M.L.J. Hautus, A polynomial characterization of (A, B)-invariant and reachability subspaces, SIAM J. Control. Optim. 18 (1980) 420–436.
[2] G.D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, SIAM J. Control. Optim. 13 (1973) 493–520.
[3] P.A. Fuhrmann, Algebraic system theory: an analyst's point of view, J. Franklin Inst. 301 (1976) 521–540.
[4] P.A. Fuhrmann, On strict system equivalence and similarity, Int. J. Contr. 25 (1977) 5–10.
[5] P.A. Fuhrmann, Linear feedback via polynomial models, Int. J. Contr. 30 (1979) 363–377.
[6] P.A. Fuhrmann, Linear Operators and Systems in Hilbert Space, McGraw-Hill, New York, 1981.
[7] P.A. Fuhrmann, Duality in polynomial models with some applications to geometric control theory, IEEE Trans. Automat. Control AC-26 (1981) 284–295.
[8] P.A. Fuhrmann, A Polynomial Approach to Linear Algebra, Springer-Verlag, New York, 1996.
[9] P.A. Fuhrmann, On behavior homomorphisms and system equivalence, Syst. Control Lett. 44 (2001) 127–134.
[10] P.A. Fuhrmann, A study of behaviors, Linear Algebra Appl. 351–352 (2002) 303–380.
[11] P.A. Fuhrmann, A note on continuous behavior homomorphisms, Syst. Control Lett. 49 (2003) 359–363.
[12] P.A. Fuhrmann, U. Helmke, On the parametrization of conditioned invariant subspaces and observer theory, Linear Algebra Appl. 332–334 (2001) 265–353.
[13] P.A. Fuhrmann, J.C. Willems, A study of (A, B)-invariant subspaces via polynomial models, Int. J. Contr. 31 (1980) 467–494.
[14] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press, 1982.
[15] M.L.J. Hautus, M. Heymann, Feedback – an algebraic approach, SIAM J. Contr. 16 (1978) 83–105.
[16] D. Hinrichsen, D. Prätzel-Wolters, Solution modules and system equivalence, Int. J. Contr. 32 (1980) 777–802.
[17] R.E. Kalman, Lectures on Controllability and Observability, CIME Summer Course, Cremonese, Roma, 1968.
[18] R.E. Kalman, P. Falb, M. Arbib, Topics in Mathematical System Theory, McGraw-Hill, 1969.
[19] M. Kuijper, First order Representations of Linear Systems, Birkhauser Verlag, 1994.
[20] J.W. Polderman, Proper elimination of latent variables, in: Proc. 12th World Congress IFAC, 1992, pp. 73–76.
[21] P. Rapisarda, J.C. Willems, State maps for linear systems, SIAM J. Contr. Optim. 35 (1997) 1053–1091.
[22] H.H. Rosenbrock, State-space and Multivariable Theory, John Wiley, New York, 1970.
[23] J. Rosenthal, J.M. Schumacher, Realization by inspection, IEEE Trans. Automat. Control 42 (1997) 1257–1263.
[24] H. Weber, Lehrbuch der Algebra, reprinted by Chelsea, New York, 1898.
[25] J.C. Willems, From time series to linear systems. Part I: Finite-dimensional linear time invariant systems, Automatica 22 (1986) 561–580.

[26] J.C. Willems, Models for dynamics, in: U. Kirchgraber, H.O. Walther (Eds.), Dynamics Reported, vol. 2, Wiley-Teubner, 1989, pp. 171–269.
[27] J.C. Willems, Paradigms and puzzles in the theory of dynamical systems, IEEE Trans. Automat. Control AC-36 (1991) 259–294.
[28] J.C. Willems, On interconnections, control and feedback, IEEE Trans. Automat. Control AC- 42 (1997) 326–337.
[29] W.A. Wolovich, Linear Multivariable Systems, Springer-Verlag, New York, 1974.